

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (Other than US) AMRAD OPERATIONS PTY LTD
(US Only)

(ii) TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS

(iii) NUMBER OF SEQUENCES: 49

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE
(B) STREET: 1 LITTLE COLLINS STREET
(C) CITY: MELBOURNE
(D) STATE: VICTORIA
(E) COUNTRY: AUSTRALIA
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT INTERNATIONAL
(B) FILING DATE: 31-OCT-1997

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: P05117
(B) FILING DATE: 14-FEB-1997

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PO 3384
(B) FILING DATE: 01-NOV-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES DR, E JOHN L
(C) REFERENCE/DOCKET NUMBER: EJH/EK

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777
(B) TELEFAX: +61 3 9254 2770

SEARCHED : 05/06/2000
SERIALIZED : 10/04/2000

- 108 -

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGCCGGCCC ACGTGAAAGGC

20

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTCGCCAATG ACAAGACGCT

20

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1236 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..636

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGAGGGCTCAA GCTCCGGGCG GATTCTGCGT GCCGCTCTCG CTCCCTGGGG TCTGTTGGCC

-101

GGCCTGTGCC ACCCGGACGC CGGGCTCACT GCCTCTGTCT CCCCATCAG CGCAGCCCCG

-41

GACGCTATGG CCCACCCCTC CAGCTGGCCC CTCGAGTAGG

-1

ATG GTA GCA CGC AAC CAG GTG GCA GCC GAC AAT GCG ATC TCC CCG GCA
Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala
1 5 10 15

48

GCA GAG CCC CGA CGG CGG TCA GAG CCC TCC TCG TCC TCG TCT TCG TCC
Ala Glu Pro Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser Ser Ser
20 25 30

96

TCG CCA GCG GCC CCC GTG CGT CCC CGG CCC TGC CCG GCG GTC CCA GCC

144

PROPERIEHYSOC1 PRV.3/10/97

- 109 -

Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala	
35 40 45	
CCA GCC CCT GGC GAC ACT CAC TTC CGC ACC TTC CGC TCC CAC TCC GAT	192
Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp	
50 55 60	
TAC CGG CCC ATC ACG CGG ACC AGC GCG CTC CTG GAC GCC TGC GGC TTC	240
Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe	
65 70 75 80	
TAT TGG GGA CCC CTG AGC GTG CAC GGG GCG CAC GAG CGG CTG CGT GCC	288
Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala	
85 90 95	
GAG CCC GTG GGC ACC TTC TTG GTG CGC GAC AGT CGT CAA CGG AAC TGC	336
Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys	
100 105 110	
TTC TTC GCG CTC AGC GTG AAG ATG GCT TCG GGC CCC ACG AGC ATC CGC	384
Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg	
115 120 125	
G TG CAC TTC CAG GCC GGC CGC TTC CAC TTG GAC GGC AGC CGC GAG ACC	432
Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr	
130 135 140	
TTC GAC TGC CTT TTC GAG CTG CTG GAG CAC TAC GTG GCG GCG CCG CGC	480
Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg	
145 150 155 160	
CGC ATG TTG GGG GCC CCG CTG CGC CAG CGC CGC GTG CGG CCG CTG CAG	528
Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln	
165 170 175	
GAG CTG TGT CGC CAG CGC ATC GTG GCC GCC GTG GGT CGC GAG AAC CTG	576
Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu	
180 185 190	
GCG CGC ATC CCT CTT AAC CCG GTA CTC CGT GAC TAC CTG AGT TCC TTC	624
Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe	
195 200 205	
CCC TTC CAG ATC TGA CCGGCTG CCGCTGTGCC GCAGCATTAA GTGGGGCCGC	676
Pro Phe Gln Ile *	
210	
CTTATTATTT CTTATTATTA ATTATTATTA TTTTCTGGA ACCACGTGGG AGCCCTCCCC	736
GCCTGGCTCG GAGGGAGTGG TTGTGGAGGG TGAGATGCCT CCCACTTCTG GCTGGAGACC	796
TCATCCCACC TCTCAGGGGT GGGGGTGCTC CCCTCCTGGT GCTCCCTCCG GGTCCCCCT	856
GGTTGTAGCA GCTTGTGTCT GGGGCCAGGA CCTGAATTCC ACTCCTACCT CTCCATGTTT	916
ACATATTCCC AGTATCTTG CACAAACCAG GGGTCGGGGA GGGTCTCTGG CTTCATTTTT	976
CTGCTGTGCCA GAATATCCTA TTTATATTT TTACAGCCAG TTTAGGTAAT AAACTTTATT	1036
ATGAAAGTTT TTTTTAAAAA GAAAAAAA AAAAAAAA	1075

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 212 amino acids

- 110 -

- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala
 1 5 10 15

Ala Glu Pro Arg Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser
 20 25 30

Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala
 35 40 45

Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp
 50 55 60

Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe
 65 70 75 80

Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala
 85 90 95

Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys
 100 105 110

Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg
 115 120 125

Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr
 130 135 140

Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg
 145 150 155 160

Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln
 165 170 175

Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu
 180 185 190

Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe
 195 200 205

Pro Phe Gln Ile
 210

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1121 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 223..819

- 111 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGATCTGTC	GGTGACAGTG	TCTGCGAGAG	ACTTTGCCAC	ACCATTCTGC	CGGAATTGG	60
AGAAAAAGAA	CCAGCCGCTT	CCAGTCCCCCT	CCCCCTCCGC	CACCATTTCG	GACACCCCTGC	120
ACACTCTCGT	TTTGGGGTAC	CCTGTGACTT	CCAGGCAGCA	CGCGAGGTCC	ACTGGCCCCA	180
GCTCGGGCGA	CCAGCTGTCT	GGGACGTGTT	GACTCATCTC	CC ATG ACC CTG CGG		234
				Met Thr Leu Arg		
				1		
TGC CTG GAG CCC TCC	GGG AAT GGA GCG GAC	AGG ACG CGG AGC CAG TGG				282
Cys Leu Glu Pro Ser	Gly Asn Gly Ala Asp	Arg Thr Arg Ser Gln Trp				
5	10	15	20			
GGG ACC GCG GGG TTG	CCG GAG GAA CAG TCC CCC	GAG GCG GCG CGT CTG				330
Gly Thr Ala Gly Leu	Pro Glu Gln Ser Pro	Glu Ala Ala Arg Leu				
25	30	35				
GCG AAA GCC CTG CGC	GAG CTC AGT CAA ACA GGA TGG TAC	TGG GGA AGT				378
Ala Lys Ala Leu Arg	Glu Leu Ser Gln Thr	Gly Trp Tyr Trp Gly Ser				
40	45	50				
ATG ACT GTT AAT GAA	GCC AAA GAG AAA TTA AAA GAG GCT	CCA GAA GGA				426
Met Thr Val Asn Glu	Ala Lys Glu Lys Leu Lys Glu Ala Pro Glu Gly					
55	60	65				
ACT TTC TTG ATT AGA	GAT AGT TCG CAT TCA GAC TAC CTA CTA	ACT ATA				474
Thr Phe Leu Ile Arg	Asp Ser Ser His Ser Asp Tyr Leu Leu Thr Ile					
70	75	80				
TCC GTT AAG ACG TCA	GCT GGA CCG ACT AAC CTG CGG ATT GAG TAC CAA					522
Ser Val Lys Thr Ser	Ala Gly Pro Thr Asn Leu Arg Ile Glu Tyr Gln					
85	90	95	100			
GAT GGG AAA TTC AGA	TTG GAT TCT ATC ATA TGT GTC AAG TCC AAG CTT					570
Asp Gly Lys Phe Arg	Leu Asp Ser Ile Ile Cys Val Lys Ser Lys Leu					
105	110	115				
AAA CAG TTT GAC AGT	GTG GTT CAT CTG ATT GAC TAC TAT GTC CAG ATG					618
Lys Gln Phe Asp Ser	Val Val His Leu Ile Asp Tyr Tyr Val Gln Met					
120	125	130				
TGC AAG GAT AAA CGG	ACA GGC CCA GAA GCC CCA CGG AAT GGG ACT GTT					666
Cys Lys Asp Lys Arg	Thr Gly Pro Glu Ala Pro Arg Asn Gly Thr Val					
135	140	145				
CAC CTG TAC CTG ACC	AAA CCT CTG TAT ACA TCA GCA CCC ACT CTG CAG					714
His Leu Tyr Leu Thr	Lys Pro Leu Tyr Thr Ser Ala Pro Thr Leu Gln					
150	155	160				
CAT TTC TGT CGA CTC	GCC ATT AAC AAA TGT ACC GGT ACG ATC TGG GGA					762
His Phe Cys Arg Leu	Ala Ile Asn Lys Cys Thr Gly Thr Ile Trp Gly					
165	170	175	180			
CTG CCT TTA CCA ACA	AGA CTA AAA GAT TAC TTG GAA GAA TAT AAA TTC					810
Leu Pro Leu Pro Thr	Arg Leu Lys Asp Tyr Leu Glu Glu Tyr Lys Phe					
185	190	195				
CAG GTA TAAGTATTTC	TCTCTCTTTT TCGTTTTTTT TTAAAAAAA AAAAACACAT					866
Gln Val						
GCCTCATATA GACTATCTCC	GAATGCAGCT ATGTGAAAGA GAACCCAGAG GCCCTCTCT					926
GGATAACTGC GCAGAATTCT	CTCTTAAGGA CAGTTGGCT CAGTCTAACT TAAAGGTGTG					986

00000000000000000000000000000000

- 112 -

AAGATGTAGC	TAGGTATTTC	AAAGTTCCCC	TTAGGTAGTT	TTAGCTGAAT	GATGCTTCT	1046
TTCCTATGGC	TGCTCAAGAT	CAAATGGCCC	TTTTAAATGA	AACAAAACAA	AACAAAACAA	1106
AAAAAAAAAA	AAAAAA					1121

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 198 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Thr Leu Arg Cys Leu Glu Pro Ser Gly Asn Gly Ala Asp Arg Thr
 1 5 10 15
 Arg Ser Gln Trp Gly Thr Ala Gly Leu Pro Glu Glu Gln Ser Pro Glu
 20 25 30
 Ala Ala Arg Leu Ala Lys Ala Leu Arg Glu Leu Ser Gln Thr Gly Trp
 35 40 45
 Tyr Trp Gly Ser Met Thr Val Asn Glu Ala Lys Glu Lys Leu Lys Glu
 50 55 60
 Ala Pro Glu Gly Thr Phe Leu Ile Arg Asp Ser Ser His Ser Asp Tyr
 65 70 75 80
 Leu Leu Thr Ile Ser Val Lys Thr Ser Ala Gly Pro Thr Asn Leu Arg
 85 90 95
 Ile Glu Tyr Gln Asp Gly Lys Phe Arg Leu Asp Ser Ile Ile Cys Val
 100 105 110
 Lys Ser Lys Leu Lys Gln Phe Asp Ser Val Val His Leu Ile Asp Tyr
 115 120 125
 Tyr Val Gln Met Cys Lys Asp Lys Arg Thr Gly Pro Glu Ala Pro Arg
 130 135 140
 Asn Gly Thr Val His Leu Tyr Leu Thr Lys Pro Leu Tyr Thr Ser Ala
 145 150 155 160
 Pro Thr Leu Gln His Phe Cys Arg Leu Ala Ile Asn Lys Cys Thr Gly
 165 170 175
 Thr Ile Trp Gly Leu Pro Leu Pro Thr Arg Leu Lys Asp Tyr Leu Glu
 180 185 190
 Glu Tyr Lys Phe Gln Val
 195

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2187 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 18..695

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCTGGCTCC	GTGCGGCC	ATG GTC ACC CAC AGC AAG TTT CCC GCC GCC GGG Met Val Thr His Ser Lys Phe Pro Ala Ala Gly 1 5 10	50
ATG AGC CGC Met Ser Arg	CCC CTG GAC ACC AGC CTG CGC CTC AAG ACC TTC AGC TCC Pro Leu Asp Thr Ser Leu Arg Leu Lys Thr Phe Ser Ser 15 20 25	98	
AAA AGC GAG Lys Ser Glu 30	TAC CAG CTG GTG GTG AAC GCC GTG CGC AAG CTG CAG GAG Tyr Gln Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu 35 40	146	
AGC GGA TTC Ser Gly Phe 45	TAC TGG AGC GCC GTG ACC GGC GGC GAG GCG AAC CTG CTG Tyr Trp Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu 50 55	194	
CTC AGC GCC GAG CCC GCG GGC ACC TTT CTT ATC CGC GAC AGC TCG GAC Leu Ser Ala Glu Pro Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp 60 65 70 75	242		
CAG CGC CAC TTC ACG TTG AGC GTC AAG ACC CAG TCG GGG ACC AAG Gln Arg His Phe Thr Leu Ser Val Lys Thr Gln Ser Gly Thr Lys 80 85 90	290		
AAC CTA CGC ATC CAG TGT GAG GGG GGC AGC TTT TCG CTG CAG AGT GAC Asn Leu Arg Ile Gln Cys Glu Gly Ser Phe Ser Leu Gln Ser Asp 95 100 105	338		
CCC CGA AGC ACG CAG CCA GTT CCC CGC TTC GAC TGT GTA CTC AAG CTG Pro Arg Ser Thr Gln Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu 110 115 120	386		
GTG CAC CAC TAC ATG CCG CCT CCA GGG ACC CCC TCC TTT TCT TTG CCA Val His His Tyr Met Pro Pro Pro Gly Thr Pro Ser Phe Ser Leu Pro 125 130 135	434		
CCC ACG GAA CCC TCG TCC GAA GTT CCG GAG CAG CCA CCT GCC CAG GCA Pro Thr Glu Pro Ser Ser Glu Val Pro Glu Gln Pro Pro Ala Gln Ala 140 145 150 155	482		
CTC CCC GGG AGT ACC CCC AAG AGA GCT TAC TAC ATC TAT TCT GGG GGC Leu Pro Gly Ser Thr Pro Lys Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly 160 165 170	530		
GAG AAG ATT CCG CTG GTA CTG AGC CGA CCT CTC TCC TCC AAC GTG GCC Glu Lys Ile Pro Leu Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala 175 180 185	578		
ACC CTC CAG CAT CTT TGT CGG AAG ACT GTC AAC GGC CAC CTG GAC TCC Thr Leu Gln His Leu Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser 190 195 200	626		
TAT GAG AAA GTG ACC CAG CTG CCT GGA CCC ATT CGG GAG TTC CTG GAT Tyr Glu Lys Val Thr Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp 205 210 215	674		
CAG TAT GAT GCT CCA CTT TAAGGAGCAA AAGGGTCAGA GGGGGGCCTG Gln Tyr Asp Ala Pro Leu 220 225	722		

- 114 -

GGTCGGTCGG	TCGCCTCTCC	TCCGAGGCAC	ATGGCACAAG	CACAAAATC	CAGCCCCAAC	782
GGTCGGTAGC	TCCCAGTGAG	CCAGGGGCAG	ATTGGCTCT	TCCTCAGGCC	CTCCACTCCC	842
GCAGAGTAGA	GCTGGCAGGA	CCTGGAATTC	GTCTGAGGGG	AGGGGGAGCT	GCCACCTGCT	902
TTCCCCCTTC	CCCCAGCTCC	ACCTTCTTTC	AACTGGAGCC	AGCCGGCCTG	GCCTGGTGGG	962
ACAATAACCTT	TGACAAGCGG	ACTCTCCCT	CCCCCTCCCTC	CACACCCCT	CTGCTTCCCA	1022
AGGGAGGTGG	GGACACCTCC	AAGTGTGAA	CTTAGAACTG	CAAGGGGAAT	CTTCAAACTT	1082
TCCCCTGGA	ACTTGTGATTC	GCTTTGATTT	GGTTTGATCA	AGAGCAGGCA	CCTGGGGGAA	1142
GGATGGAAGA	GAAAAGGGTG	TGTGAAGGGT	TTTTATGCTG	GCCAAAGAAA	TAACCACTCC	1202
CACTGCCCAA	CCTAGGTGAG	GAGTGGTGGC	TCCTGGCTCT	GGGGAGAGTG	GCAAGGGGTG	1262
ACCTGAAGAG	AGCTATACTG	GTGCCAGGCT	CCTCTCCATG	GGGCAGCTAA	TGAAACCTCG	1322
CAGATCCCTT	GCACCCAGA	ACCCCTCCCG	TTGTGAAGAG	GCAGTAGCAT	TTAGAAGGGA	1382
GACAGATGAG	GCTGGTGAGC	TGGCCGCCCT	TTCCAACACC	GAAGGGAGGC	AGATCAACAG	1442
ATGAGCCATC	TTGGAGCCA	GGTTTCCCT	GGAGCAGATG	GAGGGTTCTG	CTTGTCTCT	1502
CCTATGTGGG	GCTAGGAGAC	TCGCCTTAAA	TGCCCTCTGT	CCCAGGGATG	GGGATTGGCA	1562
CACAAGGAGC	CAAACACAGC	CAATAGGCAG	AGAGTTGAGG	GATTCAACCA	GGTGGCTACA	1622
GGCCAGGGGA	AGTGGCTGCA	GGGGAGAGAC	CCAGTCACTC	CAGGAGACTC	CTGAGTTAAC	1682
ACTGGGAAGA	CATTGGCCAG	TCCTAGTCAT	CTCTCGGTCA	GTAGGTCCGA	GAGCTTCCAG	1742
GCCCTGCACA	GCCCTCTTT	CTCACCTGGG	GGGAGGCAGG	AGGTGATGGA	GAAGCCTTCC	1802
CATGCCGCTC	ACAGGGCCCT	CACGGGAATG	CAGCAGCCAT	GCAATTACCT	GGAACCTGGTC	1862
CTGTGTTGGG	GAGAAACAAG	TTTCTGAAG	TCAGGTATGG	GGCTGGGTGG	GGCAGCTGTG	1922
TGTTGGGGTG	GCTTTTTCT	CTCTGTTTG	AATAATGTT	ACAATTGCC	TCAATCACTT	1982
TTATACTCAAT	CCACCTCCAG	CCCCCCCCCTC	TCCCCACTCA	GGCCTTCGAG	GCTGTCTGAA	2042
GATGCTTGAA	AAACTCAACC	AAATCCAGT	TCAACTCAGA	CTTTGCACAT	ATATTTATAT	2102
TTATACTCAG	AAAACACACA	TTTCAGTAAT	TTATAATAA	AGACCACTAT	TTTTTAATGA	2162
AAAAAAAAAA	AAAAAA	AAAAAA	AAAAA			2187

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 225 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Thr His Ser	Lys Phe Pro Ala Ala Gly	Met Ser Arg Pro Leu	
1	5	10	15
Asp Thr Ser Leu Arg Leu Lys	Thr Phe Ser Ser Lys Ser Glu Tyr Gln		
20	25	30	

- 115 -

Leu	Val	Val	Asn	Ala	Val	Arg	Lys	Leu	Gln	Glu	Ser	Gly	Phe	Tyr	Trp
			35				40							45	
Ser	Ala	Val	Thr	Gly	Gly	Glu	Ala	Asn	Leu	Leu	Ser	Ala	Glu	Pro	
			50			55							60		
Ala	Gly	Thr	Phe	Leu	Ile	Arg	Asp	Ser	Ser	Asp	Gln	Arg	His	Phe	Phe
			65		70			75						80	
Thr	Leu	Ser	Val	Lys	Thr	Gln	Ser	Gly	Thr	Lys	Asn	Leu	Arg	Ile	Gln
				85				90						95	
Cys	Glu	Gly	Gly	Ser	Phe	Ser	Leu	Gln	Ser	Asp	Pro	Arg	Ser	Thr	Gln
				100				105					110		
Pro	Val	Pro	Arg	Phe	Asp	Cys	Val	Leu	Lys	Leu	Val	His	His	Tyr	Met
			115				120						125		
Pro	Pro	Pro	Gly	Thr	Pro	Ser	Phe	Ser	Leu	Pro	Pro	Thr	Glu	Pro	Ser
			130			135						140			
Ser	Glu	Val	Pro	Glu	Gln	Pro	Pro	Ala	Gln	Ala	Leu	Pro	Gly	Ser	Thr
			145			150			155				160		
Pro	Lys	Arg	Ala	Tyr	Tyr	Ile	Tyr	Ser	Gly	Gly	Glu	Lys	Ile	Pro	Leu
				165				170					175		
Val	Leu	Ser	Arg	Pro	Leu	Ser	Ser	Asn	Val	Ala	Thr	Leu	Gln	His	Leu
			180				185						190		
Cys	Arg	Lys	Thr	Val	Asn	Gly	His	Leu	Asp	Ser	Tyr	Glu	Lys	Val	Thr
			195				200					205			
Gln	Leu	Pro	Gly	Pro	Ile	Arg	Glu	Phe	Leu	Asp	Gln	Tyr	Asp	Ala	Pro
			210			215						220			
Leu															
			225												

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1094 base pairs
 - (B) TYPE: nucleic acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCCGGCTGG	CCCCTTCTGT	AGGATGGTAG	CACACAACCA	GGTGGCAGCC	GACAATGCAG	60
TCTCCACAGC	AGCAGAGCCC	CGACGGCGGC	CAGAACCTTC	CTCCTCTTCC	TCCTCCTCGC	120
CCGGGGCCCC	CGGGCCCCCG	CGGCCGTGCC	CCGGCGTCCC	GGCCCCGGCC	CCCGGGGACA	180
CGCACTTCCG	CACATTCCGT	TCGCACGGCG	ATTACCGGCG	CATCACGCGC	GCCAGCGCGC	240
TCCTGGACGC	CTGCGGATTC	TACTGGGGGC	CCCTGAGCGT	GCACGGGGCG	CACGAGCGGC	300
TGCGCGCCGA	GCCCCTGGGC	ACCTTCTGG	TGCGCGACAG	CCGCCAGCGG	AACTGCTTT	360
TCGCCCTTAG	CGTGAAAGATG	GCCTCGGGAC	CCACGAGCAT	CCGCCTGCAC	TTTCAGGCCG	420
GCCGCTTCA	CCTGGATGGC	AGCCGCGAGA	GCTTCGACTG	CCTCTTCGAG	CTGCTGGAGC	480

ACTACGTGGC	GGGGCCGCGC	CGCATGCTGG	GGGCCCCGCT	GCGCCAGCGC	CCGCGTGC	540		
CGCTGCAGGA	GCTGTGCCGC	CAGCGCATCG	TGGCCACCGT	GGGCGCGAG	AACCTGGCTC	600		
GCATCCCCCT	CAAA	CCCCGTC	CTCCCGGACT	ACCTGAGCTC	CTTCCCCCTTC	CAGATTGAC	660	
CGGCAGCGCC	CGCCCGTGCAC	GCAGCATTAA	CTGGGATGCC	GTGTTATTTT	TTTATTACTT	720		
GCCTGGAACC	ATGTGGGTAC	CCTCCCCGGC	CTGGGTTGGA	GGGAGCGGAT	GGGTGTAGGG	780		
GCGAGGC	GGCC	CCCCCTC	GGCTGGAGAC	GAGGCCGCAG	ACCCCTCTC	ACCTCTTGAG	840	
GGGGTCC	CCCTCCTGGT	GCTCCCCTCG	GGTCCCCCTG	GTTGTTGTAG	CAGCTTA	900		
GTATCTGGAG	CCAGGACCTG	AACTCGCAC	TCCTACCTCT	TCATGTTTAC	ATATA	ACCCAG	960	
TATCTTGCA	CAAAC	CAAGGG	GTTGGGGAG	GGTCTCTGGC	TTTATT	TTTC	TGCTGTGCAG	1020
AATCCTATTT	TATAT	TTTTT	AAAGTCAGTT	TAGGTAATAA	ACTTTATTAT	GAAAGTTTTT	1080	
TTTTTTAAAAA	AAAA						1094	

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Val Ala His Asn Gln Val Ala Ala Asp Asn Ala Val Ser Thr Ala			
<u>1</u>	5	10	15
Ala Glu Pro Arg Arg Arg Pro Glu Pro Ser Ser Ser Ser Ser Ser Ser			
20	25	30	
Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Ala Val Pro Ala Pro			
35	40	45	
Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ala Asp Tyr			
50	55	60	
Arg Arg Ile Thr Arg Ala Ser Ala Leu Leu Asp Ala Cys Gly Phe Tyr			
65	70	75	80
Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala Glu			
85	90	95	

- 117 -

Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys Phe
100 105 110

Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg Val
115 120 125

His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Ser Phe
130 135 140

Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg Arg
145 150 155 160

Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln Glu
165 170 175

Leu Cys Arg Gln Arg Ile Val Ala Thr Val Gly Arg Glu Asn Leu Ala
180 185 190

Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe Pro
195 200 205

Phe Gln Ile
210

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2807 base pairs
 - (B) TYPE: nucleic acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAAACCGAG	CCGGGGAGAC	CAGGAGGCCT	TGGCCTCAGA	GCTTCAGAGT	CGCGTGGCAG	60
CAAACAGAGA	AACCTGTAGA	GGGCAGTGTG	CGTCACTTAG	CTCAGGGAAAG	CTGCCACGCGA	120
AACTCACCCG	CCTTCATTCA	TAAAACATCGT	CAGCTAGGCA	CCTACTCCTG	GGCTTCAGG	180
ACAAACTGAA	TCACGAAACC	ACAGTGTCT	TAAAATAGGT	CTGACCGCCT	GAATCCCTGG	240
CCAAGGTGTG	TACGGGGCAT	GGGAGCCCTT	GTGCAGAGAT	GCTTGCAGGA	GCCTTGAGGG	300
GCTCTGTAAG	ACAGAGGCTA	GGAAAGACAAA	GTTGGGGGCT	ACAGCTCTT	GTCCTGCCCG	360
GGGCTCTAGT	TTCTTCGGTT	GCCCCACGTAG	GAGTGCAGAG	AGTCCAGCCC	CTGGGGACCC	420
AACCCCAACCC	CGCCCAGTTT	CCGAGGAAC	CGTCCGGGAG	CGGGGGCGCC	CCTCCCGCAC	480
CGCCCTTAGGC	TTCCCTTGAA	GCCTCTGCGG	TCAGGCCACC	GCTTCCTGGG	AAGCCCAAGC	540

CAAGGCCAGG CCGACTGGCC AACGGGAGGG GCCCCGGCCG GATTCTGGAG GAGGGCGCG	600
GCCCCACAGG TCTCCAGGC TGGCTAGCCG GGCTCCTAGA CGGGAGACTG CCAAGGCCTT	660
CGGGTCTGG GCAGGAACGA TCCTGGCAGG GAGGAGTTGC TTGGGGGTG GGGGGAAAG	720
GCTCCAGGG CGGTGGAGCT CTGACCAGGA GAATGCACAC ACTCGGAGGG GAGGAGGCCT	780
GTCAGCCCCA AGCTAGCATC CCACCCGGGG AGCAGCGATG TGGGCGAAG GTAGCCAGAG	840
CAAAAGAGCA GGCACCCAGGT GACACGAAAC AGAAGATTC GGGTAGAGCC AGAACCCCAG	900
AAGTCCCATT CAGGGAAGGT GCGAGGCAGAG AACGAGTTAG GTGGACCCCTC TCCAGGGCAG	960
GCCAAAGAAA TCTAAAGAGA ACCCGAAGGA CTTGCCGAA AGAGAAACCG AAAGCGGCAG	1020
TGGGCGGGAT CGGTGGCGG GCCCTCCCTG GTTTAAGAGC TTGATGCAGG GGCAGGGCAGC	1080
AGCAGAGAGA ACTGCGCCG TGGCAGCGC ACGGCTCCCG GCCCCGGAGC ATGCGCGACA	1140
GCAGCCCCCG AACCCCCAGC CGCGCGCC CGCGTCCCG CGCCAGGTGA GCCGAGGCAG	1200
CTGCGAAGGA GCAGGGGGAGA GGGGATGGGA GGAAGGGGAG CAGAGCCTGG CAGGACTATC	1260
CTCGCAGACT GCATGGCGGG GTCTGGATG CTATGCCTCT GGCGCCCGCC CCACCGGCTG	1320
GCCCAGGCGG CCCCTCGCGC GCGCGGGCG CGCTCAGCCC CTCCCTCTCCG GCCCTGAGCC	1380
CGGATCGTCC GCCCAGGTTC CAGTTCCCG CGTGGCCAGT AGGCGGCAAC CGCGAGGCAG	1440
CAAGCCACCC AGCGGGGAGC GCCTGGACTC GGGCCCTCT CCACGCCCGG TTCTCCACGC	1500
CGCGGGGAG GCAGGGCTCC ACCGCCAGTC TGGAAAGGTT CCACATACAG GAACGGCTA	1560
CTTCGCGAGAT GAGCCCACCG AGGCTCAGGC TCCGGGGAGA TTCTGCGTGT CACCCCTCGCT	1620
CCTTGGGCTC CGCTGGCGG CCTGTGCCAC CGGGACGCC CGTTCACTGC CTCTGCTCC	1680
CCCATCAGCG CAGCCCCGGA CGCTATGGCC CACCCCTCCA GCTGGCCCT CGAGTAGGAT	1740
GGTAGCACGT AACCAGGTGG AAGCCGACAA TGCGATCTCC CGGGCATCAG AGCCCCGACG	1800
CGGGCCAGAG CCATCCTCGT CCTCGTCTTC GTCTCGCCG GCGGGCCCGG CGCGTCCCCG	1860
GCCCTGCCCC GTGGTCCCCGG CCCCCGGCTCC GGGCGACACT CACTTCCGCA CCTTCCGCTC	1920
CCACTCTGAT TACCGCGCA TCACCGGGAC CAGCGCTCTC CTGGACGCC GCGGCTTCTA	1980
CTGGGGACCC CTGAGCGTGC ATGGGGCGCA CGAACGGCTG CGTTCGAAC CGTGGGCCAC	2040
CTTCTTGGTG CGCGACAGTC GCCAGCGGAA CTGCTTCTTC GCGCTCAGCG TGAAGATGGC	2100
TTCGGGCCCC ACGAGCATTG GTGTGCACTT CCAGGCGGC CGCTTCCACC TGGACGGCAA	2160
CCCGGAGACC TTCGACTGCC TCTTCGAGCT GCTGCAGCAC TACGTGGCGG CGCCCGCGCCG	2220
CATGTTGGGG GCCCCACTGC GCCAGCGCCG CGTGGGGCG CTGCAAGGAGC TGTGTGCCA	2280
GGCGATCGTG CCCGCCGTGG GTCCGGAGAA CCTGGCACGC ATCCCTCTTA ACCCCGGTACT	2340
CCGTGACTAC CTGAGTTCTT TCCCCCTCCA GATCTGACCG GCTGCCGCC GCGCCGCAGA	2400
ATTAAGTGGG AGCGCCTTAT TATTCTTAT TATTAATTAT TATTATTTT CTGGAACCAC	2460
GTGGGAGGCC TCCCCGCCA GGTCGGAGGG AGTGGGTGTG GAGGGTGAGA TCCCTCCAC	2520
TTCTGGCTGG AGACCTTATC CGCCCTCTCG GGGGGCCTCC CCTCTGGTG CTCCCTCCCC	2580
GTCCCCCTGG TTGTAGCAGC TTGTGTCTGG GGCCAGGACC TGAACCTCAC GCCTACCTCT	2640
CCATGTTTAC ATGTTCCCAG TATCTTGCA CAAACCAAGGG GTGGGGAGG GTCTCTGGCT	2700
TCATTTTCT GCTGTGCAGA ATATTCTATT TTATATTTT ACATCCAGTT TAGATAATAA	2760
ACTTTATTAT GAAAGTTTTT TTTTTAAAG AAACAAAGAT TTCTAGA	2807

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 212 amino acids

- (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Glu Pro Arg Arg Arg Pro Glu Pro Ser Ser Ser Ser Ser Ser Ser
20 25 30

Ser Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Val Val Val Pro Ala
35 40 45

Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp
50 55 60

Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe
 65 70 75 80

Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ser
85 90 95

-Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys
100 105 110

Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg
115 120 125

Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Asn Arg Glu Thr
130 135 140

Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg
 145 150 155 160

Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln
165 170 175

Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu
180 185 190

- 120 -

Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe		
195	200	205
Pro Phe Gln Ile		
210		

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1611 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 263..1529

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGAATTCCGG GCGGGCTGTG TGAGTCTGTG AGTGGAAAGGC GCGCCGGCTC TTTTGTCTGA	60
GTTGTACCCCG GTGGCTTTGT TCCAGGCATT CCGGTGATTT CCTCCGGGCA GTCCGCAGAA	120

GCCGCAGCGG CCGCCCCGCGC TCTCTCTGCA GTCTCCACAC CCGGGAGAGC CTGAGCCCGC	180
GTCACGCCCG TCAGCCCCCG CTGAGTCCCT TCTCTGTTGT CGCGTCCGAA TCGAGTTCCC	240
GGAATCAGAC GGTGCCCAT AG ATG GCC AGC TTT CCC CCG AGG GTT AAC GAG	292
Met Ala Ser Phe Pro Pro Arg Val Asn Glu	
1 5 10	
AAA GAG ATC GTG AGA TCA CGT ACT ATA GGG GAA CTC TTG GCT CCA GCA	340
Lys Glu Ile Val Arg Ser Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala	
15 20 25	
GCT CCT TTT GAC AAG AAA TGT GGT GAG AAC TGG ACG GTT GCT TTT	388
Ala Pro Phe Asp Lys Lys Cys Gly Gly Glu Asn Trp Thr Val Ala Phe	
30 35 40	

- 121 -

GCT CCT GAT GGT TCC TAC TTT GCG TGG TCA CAA GGA TAT CGC ATA GTG Ala Pro Asp Gly Ser Tyr Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val	436
45 50 55	
AAG CTT GTC CCG TGG TCC CAG TGC CGT AAG AAC TTT CTT TTG CAT CGT Lys Leu Val Pro Trp Ser Gln Cys Arg Lys Asn Phe Leu Leu His Gly	484
60 65 70	
TCC AAA AAT GTT ACC AAT TCA AGC TGT CTA AAA TTG GCA AGA CAA AAC Ser Lys Asn Val Thr Asn Ser Ser Cys Leu Lys Leu Ala Arg Gln Asn	532
75 80 85 90	
AGT AAT GGT GGT CAG AAA AAC AAG CCT CCT GAG CAC GTT ATA GAC TGT Ser Asn Gly Gly Gln Lys Asn Lys Pro Pro Glu His Val Ile Asp Cys	580
95 100 105	
GGA GAC ATA GTC TGG AGT CTT GCT TTT GGG TCT TCA GTT CCA GAA AAA Gly Asp Ile Val Trp Ser Leu Ala Phe Gly Ser Ser Val Pro Glu Lys	628
110 115 120	
CAG AGT CGT TGC GTT AAT ATA GAA TGG CAT CGG TTC CGA TTT GGA CAG Gln Ser Arg Cys Val Asn Ile Glu Trp His Arg Phe Arg Phe Gly Gln	676
125 130 135	
GAT CAG CTA CTC CTT GCC ACA GGA TTA AAC AAT GGT CGC ATC AAA ATC Asp Gln Leu Leu Ala Thr Gly Leu Asn Asn Gly Arg Ile Lys Ile	724
140 145 150	
- TGG GAT GTA TAT ACA GGA AAA CTC CTC CTT AAT TTG GTA GAC CAC ATT Trp Asp Val Tyr Thr Gly Lys Leu Leu Leu Asn Leu Val Asp His Ile	772
155 160 165 170	
GAA ATG GTT AGA GAT TTA ACT TTT GCT CCA GAT GGG AGC TTA CTC CTT Glu Met Val Arg Asp Leu Thr Phe Ala Pro Asp Gly Ser Leu Leu Leu	820
175 180 185	
GTA TCA GCT TCA AGA GAC AAA ACT CTA AGA GTG TGG GAC CTG AAA GAT Val Ser Ala Ser Arg Asp Lys Thr Leu Arg Val Trp Asp Leu Lys Asp	868
190 195 200	
GAT GGA AAC ATG GTG AAA GTA TTG CGG GCA CAT CAG AAT TGG GTG TAC Asp Gly Asn Met Val Lys Val Leu Arg Ala His Gln Asn Trp Val Tyr	916
205 210 215	
AGT TGT GCA TTC TCT CCC GAC TGT TCT ATG CTG TGT TCA GTG GGC GCC	964

- 122 -

Ser Cys Ala Phe Ser Pro Asp Cys Ser Met Leu Cys Ser Val Gly Ala			
220	225	230	
AGT AAA GCA GTT TTC CTT TGG AAT ATG GAT AAA TAC ACC ATG ATT AGG			1012
Ser Lys Ala Val Phe Leu Trp Asn Met Asp Lys Tyr Thr Met Ile Arg			
235	240	245	250
AAG CTG GAA GCT CAT CAC CAT GAT GTT GTA GCT TGT GAC TTT TCT CCT			1060
Lys Leu Glu Gly His His Asp Val Val Ala Cys Asp Phe Ser Pro			
255	260	265	
GAT GGA GCA TTG CTA GCT ACT GCA TCC TAT GAC ACT CGT GTG TAT GTC			1108
Asp Gly Ala Leu Leu Ala Thr Ala Ser Tyr Asp Thr Arg Val Tyr Val			
270	275	280	
TGG GAT CCA CAC AAT GGA GAC CTT CTG ATG GAG TTT GGG CAC CTG TTT			1156
Trp Asp Pro His Asn Gly Asp Leu Leu Met Glu Phe Gly His Leu Phe			
285	290	295	
CCC TCG CCC ACT CCA ATA TTT GCT GGA GGA GCA AAT GAC CGA TGG GTG			1204
Pro Ser Pro Thr Pro Ile Phe Ala Gly Gly Ala Asn Asp Arg Trp Val			
300	305	310	
AGA GCT GTG TCT TTC AGT CAT GAT GGA CTG CAT GTT GCC AGC CTT GCT			1252
Arg Ala Val Ser Phe Ser His Asp Gly Leu His Val Ala Ser Leu Ala			
315	320	325	330
GAT GAT AAA ATG GTG AGG TTC TGG AGA ATC GAT GAG GAT TGT CCG GTA			1300
Asp Asp Lys Met Val Arg Phe Trp Arg Ile Asp Glu Asp Cys Pro Val			
335	340	345	
CAA GTT GCA CCT TTG AGC AAT GGT CTT TGC TGT GCC TTT TCT ACT GAT			1348
Gln Val Ala Pro Leu Ser Asn Gly Leu Cys Cys Ala Phe Ser Thr Asp			
350	355	360	
GGC AGT GTT TTA GCT GCT GGG ACA CAT GAT GGA AGT GTG TAT TTT TGG			1396
Gly Ser Val Leu Ala Ala Gly Thr His Asp Gly Ser Val Tyr Phe Trp			
365	370	375	
GCC ACT CCA AGG CAA GTC CCT AGC CTT CAA CAT ATA TGT CGC ATG TCA			1444
Ala Thr Pro Arg Gln Val Pro Ser Leu Gln His Ile Cys Arg Met Ser			
380	385	390	
ATC CGA AGA GTG ATG TCC ACC CAA GAA GTC CAA AAA CTG CCT GTT CCT			1492
Ile Arg Arg Val Met Ser Thr Gln Glu Val Gln Lys Leu Pro Val Pro			

- 123 -

395	400	405	410
TCC AAA ATA TTG GCG TTT CTC TCC TAC CGC GGT TAG A CTGAAGACTG			
Ser Lys Ile Leu Ala Phe Leu Ser Tyr Arg Gly *			
415		420	
CCTTCCTGG TAGGCCTGCC AGACAGAGCG CCCTTTACAA GACACACCTC AAGCTTTACC			
TCGTGCCGAA TT			
			1539
			1599
			1611

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Ser Phe Pro Pro Arg Val Asn Glu Lys Glu Ile Val Arg Ser			
1	5	10	15
Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala Ala Pro Phe Asp Lys Lys			
20		25	30
Cys Gly Gly Glu Asn Trp Thr Val Ala Phe Ala Pro Asp Gly Ser Tyr			
35	40		45
Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val Lys Leu Val Pro Trp Ser			
50	55		60
Gln Cys Arg Lys Asn Phe Leu Leu His Gly Ser Lys Asn Val Thr Asn			
65	70	75	80
Ser Ser Cys Leu Lys Leu Ala Arg Gln Asn Ser Asn Gly Gly Gln Lys			
85		90	95
Asn Lys Pro Pro Glu His Val Ile Asp Cys Gly Asp Ile Val Trp Ser			
100	105		110
Leu Ala Phe Gly Ser Ser Val Pro Glu Lys Gln Ser Arg Cys Val Asn			
115	120		125

- 124 -

Ile Glu Trp His Arg Phe Arg Phe Gly Gln Asp Gln Leu Leu Ala
130 135 140

Thr Gly Leu Asn Asn Gly Arg Ile Lys Ile Trp Asp Val Tyr Thr Gly
145 150 155 160

Lys Leu Leu Leu Asn Leu Val Asp His Ile Glu Met Val Arg Asp Leu
165 170 175

Thr Phe Ala Pro Asp Gly Ser Leu Leu Leu Val Ser Ala Ser Arg Asp
180 185 190

Lys Thr Leu Arg Val Trp Asp Leu Lys Asp Asp Gly Asn Met Val Lys
195 200 205

Val Leu Arg Ala His Gln Asn Trp Val Tyr Ser Cys Ala Phe Ser Pro
210 215 220

Asp Cys Ser Met Leu Cys Ser Val Gly Ala Ser Lys Ala Val Phe Leu
225 230 235 240

Trp Asn Met Asp Lys Tyr Thr Met Ile Arg Lys Leu Glu Gly His His
245 250 255

His Asp Val Val Ala Cys Asp Phe Ser Pro Asp Gly Ala Leu Leu Ala
260 265 270

Thr Ala Ser Tyr Asp Thr Arg Val Tyr Val Trp Asp Pro His Asn Gly
--- 275 280 285

Asp Leu Leu Met Glu Phe Gly His Leu Phe Pro Ser Pro Thr Pro Ile
290 295 300

Phe Ala Gly Gly Ala Asn Asp Arg Trp Val Arg Ala Val Ser Phe Ser
305 310 315 320

His Asp Gly Leu His Val Ala Ser Leu Ala Asp Asp Lys Met Val Arg
325 330 335

Phe Trp Arg Ile Asp Glu Asp Cys Pro Val Gln Val Ala Pro Leu Ser
340 345 350

Asn Gly Leu Cys Cys Ala Phe Ser Thr Asp Gly Ser Val Leu Ala Ala
355 360 365

- 125 -

Gly Thr His Asp Gly Ser Val Tyr Phe Trp Ala Thr Pro Arg Gln Val
 370 375 380

Pro Ser Leu Gln His Ile Cys Arg Met Ser Ile Arg Arg Val Met Ser
 385 390 395 400

Thr Gln Glu Val Gln Lys Leu Pro Val Pro Ser Lys Ile Leu Ala Phe
 405 410 415

Leu Ser Tyr Arg Gly *
 420

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 783 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTGTCTTCCT	CCGCAGCGCG	AGGCTGGTAA	CAGGGTCTAT	TGTCTGTGGT	TGACTCCGTA	60
CTTTCGGTCTG	AGGCCTTCGG	GAGCTTTCCC	GAGGCAGTTA	GCAGAAGCCG	CAGCGACCGC	120
CCCCGCCCGT	CTCCTCTGTC	CCTGGGCCCG	GGAGACAAAC	TTGGCGTCAC	GCCCTCAGCG	180
GTCGCCACTC	TCTTCTCTGT	TGTTGGGTCC	GCATCGTATT	CCCCGAATCA	GACGGTGCCC	240
CATAGATGGC	CAGCTTTCCC	CCGAGGGTCA	ACGAGAAAGA	GATCGTGAGA	TCACGTACTA	300
TAGGTGAACT	TTTAGCTCCT	GCAGCTCCTT	TTGACAAGAA	ATGTGGTCGT	AAAAATTGGA	360
CTGTTGCTTT	TGCTCCAGAT	GGTCATACT	TTGCTTGTC	ACAAGGACAT	CGCACAGTAA	420
AGCTTGTCC	GTGGTCCCAG	TGCCTTCAGA	ACTTTCTCTT	GCATGGCACC	AAGAATGTTA	480
CCAATTCAAG	CAGTTTAAGA	TTGCCAAGAC	AAAATAGTGA	TGGTGGTCAG	AAAAATAAGC	540
CTCGTGACAT	ATTATAGACT	GTGGAGATAT	AGTCTGGAGT	CTTGCTTTG	GGTCATCAGT	600

DECODED - 15 DEC 1997

- 126 -

TCCAGAAAAA CAGAGTCGCT GTGTAAATAT AGAATGGCAT CGCTTCAGAT TTGGACAAGA	660
TCAGCTACTT CTTGCTACAG GGTTGAACAA TGGCGTATC AAAATATGGG ATGTATATCA	720
GGAAAACCTCCT CCTTAACCTTG GTAGATCATA CTGAAGTGGT CAGAGATTAA ACTTTGCTC	780
CAG	783

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1122 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTCTGTATGT CTGAATGAAG CTATAACATT TGCCTTTTA TTGCAGGTT TCCTTTGGAA	60
TATGGATAAA TACACCATGA TACGGAAACT AGAAGGACAT CACCATGATG TGGTAGCTTG	120
TGACTTTCT CCTGATGGAG CATTACTGGC TACTGCATCT TATGATACTC GAGTATATAT	180
CTGGGATCCA CATAATGGAG ACATTCTGAT GGAATTGGG CACCTGTTTC CCCCACCTAC	240
TCCAATATTT GCTGGAGGAG CAAATGACCG GTGGGTACGA TCTGTATCTT TTAGCCATGA	300
TGGACTGCAT GTTGCAGGCC TTGCTGATGA TAAAATGGTG AGGTTCTGGA GAATTGATGA	360
GGATTATCCA GTGCAAGTTG CACCTTGAG CAATGGTCTT TGCTGTGCCT TCTCTACTGA	420
TGGCAGTGTGTT TTAGCTGCTG GGACACATGA CGGAAGTGTG TATTTTGAGG CCACCTCCACG	480
GCAGGTCCCT AGCCTGCAAC ATTTATGTCG CATGTCAATC CGAAGAGTGA TGCCCACCCA	540
AGAAGTTCAAG GAGCTGCCGA TTCCCTCCAA GCTTTGGAG TTTCTCTCGT ATCGTATTTA	600
GAAGATTCTG CCTTCCCTAG TAGTAGGGAC TGACAGAATA CACTAACAC AACCTCAAG	660
CTTTACTGAC TTCAATTATC TGTTTTAAA GACGTAGAAG ATTTATTTAA TTTGATATGT	720

- 127 -

TCTTGTACTG CATTGGATC AGTTGAGCTT TTAAAATATT ATTTATAGAC AATAGAAGTA	780
TTTCTGAACA TATCAAATAT AAATTTTTTT AAAGATCTAA CTGTGAAAAC ATACATAACCT	840
GTACATATTT AGATATAAGC TGCTATATGT TGAATGGACC CTMTTGCTTT TCTGATTTTT	900
AGTTCTGACA TGTATATATT GCTTCAGTAG AGCCACAATA TGTATCTTG CTGTAAAGTG	960
CAAGGAAATT TTAAATTCTG GGACACTGAG TTAGATGGTA AATACTGACT TACGAAAGTT	1020
GAATTGGGTG AGGCAGGCAA ATCACCTGAG GTCAGCAGTT TGAGACTAGC CTGGCAAACA	1080
TGATGAAACC CTGTCTCTAC TAAAAATACA AAAAAAAAAAA AA	1122

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2537 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 422..2029

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGCACGAGC CGGGCTCCGT CCGGAGGAAG CGAGGCTGCG CCGCCGGCCC GGCAGGAGCG	60
GAGGACGGGA GCGCGGGCGG TCGCGCTCGC CCTGTCGCTG ACTGCGCTGC CCCGGCCCAT	120
CCTTGCTGG CCGCAGGTGC CCTGGATGAG GCCGCCGGCG GTGTCCCGGC CGCTGAGTGT	180
CCCCCGCGGT CGCCCGGCAGC CTGCCCTCAA GCGGCCGCCT CTCCCTGCC GGGTCCCCGT	240
TTTCCCCCGG CGCAGTCCTC CTCCGGTGGG CGCCTCCGCA CCTCGGCGCA GGCGGCACGG	300
CCCTCGGGCC GGGATGGATC CGCCGGGAAG AGGAAGACAA GCCGGGGCGT TGAGCCCCCTG	360

08962550-102197

- 128 -

CGCACGGTGC CGCCGCGCGT AGTGGGAGCT TACTCGCAGT AGGCTCTCGC TCTTCTAATC		420
A ATG GAT AAA GTG GGG AAA ATG TGG AAC AAC TTA AAA TAC AGA TGC		466
Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys		
1 5 10 15		
CAG AAT CTC TTC AGC CAC GAG GGA GGA AGC CGT AAT GAG AAC GTG GAG		514
Gln Asn Leu Phe Ser His Glu Gly Ser Arg Asn Glu Asn Val Glu		
20 25 30		
ATG AAC CCC AAC AGA TGT CCG TCT GTC AAA GAG AAA AGC ATC AGT CTG		562
Met Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu		
35 40 45		
GGA GAG GCA GCT CCC CAG CAA GAG AGC AGT CCC TTA AGA GAA AAT GTT		610
Gly Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val		
50 55 60		
GCC TTA CAG CTG GGA CTG AGC CCT TCC AAG ACC TTT TCC AGG CGG AAC		658
Ala Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn		
65 70 75		
CAA AAC TGT GCC GCA GAG ATC CCT CAA GTG GTT GAA ATC AGC ATC GAG		706
Gln Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu		
80 85 90 95		
AAA GAC AGT GAC TCG GGT GCC ACC CCA GGA ACG AGG CTT GCA CGG AGA		754
Lys Asp Ser Asp Ser Gly Ala Thr Pro Gly Thr Arg Leu Ala Arg Arg		
— 100 105 110		
GAC TCC TAC TCG CGG CAC GCC CCG TGG GGA GGA AAG AAG AAA CAT TCC		802
Asp Ser Tyr Ser Arg His Ala Pro Trp Gly Gly Lys Lys Lys His Ser		
115 120 125		
TGT TCC ACA AAG ACC CAG AGT TCA TTG GAT ACC GAG AAA AAG TTT GGT		850
Cys Ser Thr Lys Thr Gln Ser Ser Leu Asp Thr Glu Lys Lys Phe Gly		
130 135 140		
AGA ACT CGA AGC GGC CTT CAG AGG CGA GAG CGG CGC TAT GGA GTC AGC		898
Arg Thr Arg Ser Gly Leu Gln Arg Arg Glu Arg Arg Tyr Gly Val Ser		
145 150 155		
TCC ATG CAG GAC ATG GAC AGC GTT TCT AGC CGC GCG GTC GGG AGC CGC		946
Ser Met Gln Asp Met Asp Ser Val Ser Ser Arg Ala Val Gly Ser Arg		
160 165 170 175		

PROPER/EJM/SOCS1.PRV-31/10/97

- 129 -

TCC CTG AGG CAG AGG CTC CAG GAC ACG GTG GGT TTG TGT TTT CCC ATG Ser Leu Arg Gln Arg Leu Gln Asp Thr Val Gly Leu Cys Phe Pro Met	180	185	190	994	
AGA ACT TAC AGC AAG CAG TCA AAG CCA CTC TTT TCC AAT AAA AGA AAA Arg Thr Tyr Ser Lys Gln Ser Lys Pro Leu Phe Ser Asn Lys Arg Lys	195	200	205	1042	
ATA CAT CTT TCT GAA TTA ATG CTG GAG AAA TGC CCT TTT CCT GCT GGC Ile His Leu Ser Glu Leu Met Leu Glu Lys Cys Pro Phe Pro Ala Gly	210	215	220	1090	
TCG GAT TTA GCA CAA AAG TGG CAT TTG ATT AAA CAG CAT ACC GCC CCT Ser Asp Leu Ala Gln Lys Trp His Leu Ile Lys Gln His Thr Ala Pro	225	230	235	1138	
GTG AGC CCA CAC TCA ACA TTT TTT GAT ACA TTT GAT CCA TCA CTG GTG Val Ser Pro His Ser Thr Phe Phe Asp Thr Phe Asp Pro Ser Leu Val	240	245	250	255	1186
TCT ACA GAA GAT GAA GAA GAT AGG CTT CGC GAG AGA AGA CGG CTT AGT Ser Thr Glu Asp Glu Glu Asp Arg Leu Arg Glu Arg Arg Arg Leu Ser	260	265	270	1234	
ATC GAA GAA GGG GTG GAT CCC CCT CCC AAC GCA CAA ATA CAC ACC TTT Ile Glu Glu Gly Val Asp Pro Pro Asn Ala Gln Ile His Thr Phe	275	280	285	1282	
GAA GGT ACT GCA CAG GTC AAC CCA TTG TAT AAG CTG GGA CCA AAG TTA Glu Ala Thr Ala Gln Val Asn Pro Leu Tyr Lys Leu Gly Pro Lys Leu	290	295	300	1330	
GCT CCT GGG ATG ACA GAG ATA AGT GGA GAT GGT TCT GCA ATT CCA CAA Ala Pro Gly Met Thr Glu Ile Ser Gly Asp Gly Ser Ala Ile Pro Gln	305	310	315	1378	
GCA ATT GTG ACT CAG AAG AGG ATT CAA CCA CCC TAT GTC TGC AGT CAC Ala Ile Val Thr Gln Lys Arg Ile Gln Pro Pro Tyr Val Cys Ser His	320	325	330	335	1426
GGA GGC AGA AGC AGC GCC AGG TGT CCG GGG ACA GCC ACG CGC ACG TTA Gly Gly Arg Ser Ser Ala Arg Cys Pro Gly Thr Ala Thr Arg Thr Leu	340	345	350	1474	
GCA GAC AGG CAG CTT GGA AAG TTC ATA CGC AGA TCG ATT ACA TAC ACT				1522	

- 130 -

Ala Asp Arg Glu Leu Gly Lys Phe Ile Arg Arg Ser Ile Thr Tyr Thr			
355	360	365	
 GCC TCG TGC CAG ATT TGC TTC AGA TCA CAG GGA ATC CCT GTT ACT GGG			1570
Ala Ser Cys Gln Ile Cys Phe Arg Ser Gln Gly Ile Pro Val Thr Gly			
370	375	380	
 GCG TGA TGG ACC GAT ACG AGG CCG AAG CCC TTC TAG AAG GGA AAC CGG			1618
Ala * Trp Thr Asp Thr Arg Pro Lys Pro Phe * Lys Gly Asn Arg			
385	390	395	
 AAG GCA CGT TCT TGC TCA GGG ACT CTG CAC AGG AGG ACT ACC TCT TCT			1666
Lys Ala Arg Ser Cys Ser Gly Thr Leu His Arg Arg Thr Thr Ser Ser			
400	405	410	415
 CTG TGA GCT TCC GCC GCT ACA ACA GGT CTC TGC ACG CCC GGA TCG AGC			1714
Leu * Ala Ser Ala Ala Thr Thr Gly Leu Cys Thr Pro Gly Ser Ser			
420	425	430	
 AGT GGA ACC ACA ACT TCA GCT TCG ATG CCC ATG ACC CCT GCG TGT TTC			1762
Ser Gly Thr Thr Ser Ala Ser Met Pro Met Thr Pro Ala Cys Phe			
435	440	445	
 ACT CCT CCA CGT CAC GGG GCT TCT CGA ACA CTA TAA AGA CCC CAG CTC			1810
Thr Pro Pro Arg His Gly Ala Ser Arg Thr Leu * Arg Pro Gln Leu			
450	455	460	
 TTG CAT GTT TTT TGA ACC GTT GCT AAC GAT ATC ACT GAA TAG AAC TTT			1858
Leu His Val Phe * Thr Val Ala Asn Asp Ile Thr Glu * Asn Phe			
465	470	475	
 CCC TTT CAG CCT GCA GTA TAT CTG CCG CGC ACT GAT CTG CAG ATG CAC			1906
Pro Phe Gln Pro Ala Val Tyr Leu Pro Arg Ser Asp Leu Gln Met His			
480	485	490	495
 TAC GTA TGA TGG GAT TGA CGG GCT CCC GCT ACC GTC GAT GTT ACA GGA			1954
Tyr Val * Trp Asp * Arg Ala Pro Ala Thr Val Asp Val Thr Gly			
500	505	510	
 TTT TTT AAA AGA GTA TCA TTA TAA ACA AAA AGT TAG GGT TCG CTG GTT			2002
Phe Phe Lys Arg Val Ser Leu * Thr Lys Ser * Gly Ser Leu Val			
515	520	525	
 AGA ACG AGA CCA GTC AAA GCA AAG TAACTCCTGT CCCCAAAGGG CACTAACTAA			2056
Arg Thr Arg Pro Val Lys Ala Lys			

- 131 -

530

535

GTCTGCTCCT CCCGTGCATC GAACTGCACC CATAGGAGGC AGTCAGCTGC TAGGATTCC	2116
CACCCAGAAT GGGAGCTTAG TCATTAGCCT CTGCCCTATG GGGTCCCGCTG TTCCCTCAGAC	2176
AAAGGTGCCT AGGGACAGCA AGATGGCTTG CAGGTGTTCG GTGGGCTGTG ACAACTGAGG	2236
GAGGCAACTC TGGGGCATTT GCTATGAAGA ATTCTATTTC TTACCGAAGA ACAAAATTATT	2296
AATATTGGAT GGGTATTCA ATAGTGTGAC TAATGTTGA AATTATTTT TCTAAGAATT	2356
TTTCTATAAC CTTCAGAAAA AGTAGTGATG TTTGTAGTTA CTATAAATCA AGCTTTGAAA	2416
GTTCAAAACA AACAAAGTTAA ATAAAAGACT ACCTTCCTTT TAGAGAAAAC AAATGCAAGT	2476
TTTCCCAGCC ACAGGCATTG TGCACTGTTA ATGTTGCTTG TTATCAGCTC CTTTCTCCTC	2536
C	2537

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 535 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

—(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys Gln			
1	5	10	15
Asn Leu Phe Ser His Glu Gly Gly Ser Arg Asn Glu Asn Val Glu Met			
20	25	30	
Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu Gly			
35	40	45	
Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val Ala			
50	55	60	
Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn Gln			

- 132 -

65	70	75	80
Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu Lys			
85		90	95
Asp Ser Asp Ser Gly Ala Thr Pro Gly Thr Arg Leu Ala Arg Arg Asp			
100		105	110
Ser Tyr Ser Arg His Ala Pro Trp Gly Gly Lys Lys His Ser Cys			
115		120	125
Ser Thr Lys Thr Gln Ser Ser Leu Asp Thr Glu Lys Lys Phe Gly Arg			
130		135	140
Thr Arg Ser Gly Leu Gln Arg Arg Glu Arg Arg Tyr Gly Val Ser Ser			
145	150	155	160
Met Gln Asp Met Asp Ser Val Ser Ser Arg Ala Val Gly Ser Arg Ser			
165		170	175
Leu Arg Gln Arg Leu Gln Asp Thr Val Gly Leu Cys Phe Pro Met Arg			
180		185	190
Thr Tyr Ser Lys Gln Ser Lys Pro Leu Phe Ser Asn Lys Arg Lys Ile			
195		200	205
His Leu Ser Glu Leu Met Leu Glu Lys Cys Pro Phe Pro Ala Gly Ser			
210		215	220
 Asp Leu Ala Gln Lys Trp His Leu Ile Lys Gln His Thr Ala Pro Val			
225	230	235	240
Ser Pro His Ser Thr Phe Phe Asp Thr Phe Asp Pro Ser Leu Val Ser			
245		250	255
Thr Glu Asp Glu Glu Asp Arg Leu Arg Glu Arg Arg Arg Leu Ser Ile			
260		265	270
Glu Glu Gly Val Asp Pro Pro Asn Ala Gln Ile His Thr Phe Glu			
275		280	285
Ala Thr Ala Gln Val Asn Pro Leu Tyr Lys Leu Gly Pro Lys Leu Ala			
290		295	300
Pro Gly Met Thr Glu Ile Ser Gly Asp Gly Ser Ala Ile Pro Gln Ala			

305	310	315	320
Ile Val Thr Gln Lys Arg Ile Gln Pro Pro Tyr Val Cys Ser His Gly			
	325	330	335
Gly Arg Ser Ser Ala Arg Cys Pro Gly Thr Ala Thr Arg Thr Leu Ala			
	340	345	350
Asp Arg Glu Leu Gly Lys Phe Ile Arg Arg Ser Ile Thr Tyr Thr Ala			
	355	360	365
Ser Cys Gln Ile Cys Phe Arg Ser Gln Gly Ile Pro Val Thr Gly Ala			
	370	375	380
* Trp Thr Asp Thr Arg Pro Lys Pro Phe * Lys Gly Asn Arg Lys	385	390	395
	400		
Ala Arg Ser Cys Ser Gly Thr Leu His Arg Arg Thr Thr Ser Ser Leu			
	405	410	415
* Ala Ser Ala Ala Thr Thr Gly Leu Cys Thr Pro Gly Ser Ser Ser	420	425	430
Gly Thr Thr Thr Ser Ala Ser Met Pro Met Thr Pro Ala Cys Phe Thr			
	435	440	445
Pro Pro Arg His Gly Ala Ser Arg Thr Leu * Arg Pro Gln Leu Leu	450	455	460
His Val Phe * Thr Val Ala Asn Asp Ile Thr Glu * Asn Phe Pro			
	465	470	475
Phe Gln Pro Ala Val Tyr Leu Pro Arg Ser Asp Leu Gln Met His Tyr			
	485	490	495
Val * Trp Asp * Arg Ala Pro Ala Thr Val Asp Val Thr Gly Phe			
	500	505	510
Phe Lys Arg Val Ser Leu * Thr Lys Ser * Gly Ser Leu Val Arg			
	515	520	525
Thr Arg Pro Val Lys Ala Lys			
	530	535	

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1221 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATTAAACAG CATAACAGCTC CTGTGAGCCC ACATTCAACA TTTTTGATA CTTTGATCCA	60
TCTTTGGTTT CTACAGAAGA TGAAGAAGAT AGGCTTAGAG AGAGAAGGCG GCTTAGTATT	120
CAAGAAGGGG TTGATCCCC TCCCAATGCA CAAATACATA CATTGAAAGC TACTGCACAG	180
CTTAATCCAT TATTAACACTG GGACCAAAAT TAGCTCCTGG AATGACTGAA ATAAGTGGGG	240
ACAGTTCTGC AATTCCACAA GCTAATTGTG ACTCGGAAGA GGATACAACC ACCCTGTGTT	300
GCAGTCACGG AGGCAGAACG AGCGTCAGAT ATCTGGAGAC AGCCATACCC ATGTTAGCAG	360
ACAGGGAGCT TGGAAAGTCC ACACACAGAT TGATTACATA CACTGCTTCG TGCCTGATT	420
GCTT TA ATT ACAGGGAATC CCTGTTACTG GGGACTGATG GACCGTTATG AACAGAGAAC	480
CCTTCTCGAA GGGAAACCTG AAGGCACGTT TTTGCTCAGG GACTCTGCGC AAGAGGACTA	540
CTTCTCTCTCT GTGAGCTTCC GCCGATACAA CAGATCCCTG CATGCCGAA TTGAGCAGTG	600
GAATCACAAAC TTTAGTTTCG ACGCCCATGA CCCGTGTGTA TTTCACTCCT CCACGTAAAC	660
GGGACTTTA GAACATTATA AAGATCCAG TTCGTGCATG TTTTTGAAAC CATTGCTTAC	720
TATATCACTA AATAGGACTT TCCCTTTAG CCTGCAGTAT ATCTGTCGCG CGGTAATCTG	780
CAGGTGCACT ACGTATGATG GAATTGATGG GCTCCCTCTA CCCTCAATGT TACAGGATT	840
TTTAAAAGAG TATCATTATA AACAAAAAGT TAGAGTTCGC TGGTTGAAAC GAGAACCAGT	900
CAAGGCAAAG TAAACTCTCC GGTCCCCAAA GGGTGTAAAC TAGGTCCGCT TTCATGTGCA	960

- 135 -

TCAGACAGTA CACCTATAGC AACCAACACGT AGCAGTGTAA GGCTTTTCA TACAGTATGT	1020
AAGCTTAGTG TTAGTATCTG TCAGATGCTA CCTGCTGTTA CTTATTAGA TAAACATGGT	1080
GCCTATTGGA ACAATAGCGG ATAGAGCTAC AGGTGTTCAAG TAAGACTACA AAAACATTT	1140
GCCTATTTCG CTAACAGTTT GGTTTTAAT GGCTGTGGTA TTTGAGTGAG GCAACTCTGG	1200
GGCATTTGTT ATGAAGAAAT G	1221

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 116..1330

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCACGAGGC GGTGGTGGCG CGGGCGGGCG CGGCCGCGGC GGGGCGGGCG CGGAATGAAG	60	
GCCCACGGCC CTGGGGGCTG AGGCGCCCCGC CGCCTGGGGC GGGCCGCGCG TCCTC ATG	118	
	Met	
	1	
GAG GCC GGA GAG GAG CCG CTG CTG GCT GAA CTC AAG CCT GGG CGC	166	
Glu Ala Gly Glu Glu Pro Leu Leu Leu Ala Glu Leu Lys Pro Gly Arg		
5	10	15
CCC CAC CAG TTC GAC TGG AAG TCA AGC TGC GAG ACC TGG AGC GTG GCC	214	
Pro His Gln Phe Asp Trp Lys Ser Ser Cys Glu Thr Trp Ser Val Ala		
20	25	30
TTC TCG CCA GAC GGT TCC TGG TTC GCC TGG TCT CAA GGA CAC TGC GTG	262	

Phe Ser Pro Asp Gly Ser Trp Phe Ala Trp Ser Gln Gly His Cys Val
 35 40 45

GTC AAG CTG GTC CCC TGG CCC TTA GAG GAA CAG TTC ATC CCT AAA GGA 310
 Val Lys Leu Val Pro Trp Pro Leu Glu Glu Gln Phe Ile Pro Lys Gly
 50 55 60 65

TTC GAA GCC AAG AGC CGA AGC AGC AAG AAT GAC CCA AAA GGA CGG GGC 358
 Phe Glu Ala Lys Ser Arg Ser Ser Lys Asn Asp Pro Lys Gly Arg Gly
 70 75 80

AGT CTG AAG GAG AAG ACG CTG GAC TGT GGC CAG ATT GTG TGG GGG CTG 406
 Ser Leu Lys Glu Lys Thr Leu Asp Cys Gly Gln Ile Val Trp Gly Leu
 85 90 95

GCC TTC AGC CCG TGG CCC TCT CCA CCC AGC AGG AAA CTC TGG GCA CGT 454
 Ala Phe Ser Pro Trp Pro Ser Pro Ser Arg Lys Leu Trp Ala Arg
 100 105 110

CAC CAT CCC CAG GCG CCT GAT GTT TCT TGC CTG ATC CTG GCC ACA GGT 502
 His His Pro Gln Ala Pro Asp Val Ser Cys Leu Ile Leu Ala Thr Gly
 115 120 125

CTC AAC GAT GGG CAG ATC AAG ATT TGG GAG GTA CAG ACA GGC CTC CTG 550
 Leu Asn Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu Leu
 130 135 140 145

CTT CTG AAT CTT TCT GGC CAC CAA GAC GTC GTG AGA GAT CTG AGC TTC 598
 Leu Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser Phe
 150 155 160

ACG CCC AGC GGC AGT TTG ATT TTG GTC TCT GCA TCC CGG GAT AAG ACA 646
 Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys Thr
 165 170 175

CTT CGA ATT TGG GAC CTG AAT AAA CAC GGT AAG CAG ATC CAG GTG TTA 694
 Leu Arg Ile Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln Val Leu
 180 185 190

CCG GGC CAT CTG CAG TGG GTT TAC TGC TGC TCC ATC TCC CCT GAC TGT 742
 Ser Gly His Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser Pro Asp Cys
 195 200 205

GC ATG CTG TGC TCT GCA GCT GGG GAG AAG TCG GTC TTT CTG TGG AGC 790
 Ser Met Leu Cys Ser Ala Ala Gly Glu Lys Ser Val Phe Leu Trp Ser

- 137 -

210	215	220	225	
ATG CGG TCC TAC ACA CTA ATC CGG AAA CTA GAA GGC CAC CAA AGC AGT Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu Gly His Gln Ser Ser				838
230	235	240		
GTT GTC TCC TGT GAT TTC TCT CCT GAT TCA GCC TTG CTT GTC ACA GCT Val Val Ser Cys Asp Phe Ser Pro Asp Ser Ala Leu Leu Val Thr Ala				886
245	250	255		
TCG TAT GAC ACC AGT GTG ATT ATG TGG GAC CCC TAC ACC GGC GCG AGG Ser Tyr Asp Thr Ser Val Ile Met Trp Asp Pro Tyr Thr Gly Ala Arg				934
260	265	270		
CTG AGG TCA CTT CAT CAC ACA CAA CTT GAA CCC ACC ATG GAT GAC AGT Leu Arg Ser Leu His His Thr Gln Leu Glu Pro Thr Met Asp Asp Ser				982
275	280	285		
GAC GTC CAC ATG AGC TCC CTG AGG TCC GTG TGC TTC TCA CCT GAA GGC Asp Val His Met Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu Gly				1030
290	295	300	305	
TTG TAT CTC GCT ACG GTG GCA GAT GAC AGG CTG CTC AGG ATC TGG GCT Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp Ala				1078
310	315	320		
CTG GAA CTG AAG GCT CCG GTT GCC TTT GCT CCG ATG ACC AAT GGT CTT Leu Glu Leu Lys Ala Pro Val Ala Phe Ala Pro Met Thr Asn Gly Leu				1126
--	325	330	335	
TGC TGC ACG TTC TTC CCA CAC GGT GGA ATT ATT GCC ACA GGG ACG AGA Cys Cys Thr Phe Pro His Gly Gly Ile Ile Ala Thr Gly Thr Arg				1174
340	345	350		
GAT GGC CAT GTC CAG TTC TGG ACA GCT CCC CGG GTC CTG TCC TCA CTG Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser Leu				1222
355	360	365		
AAG CAC TTA TGC AGG AAA GCC CTC CGA AGT TTC CTG ACA ACG TAT CAA Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr Gln				1270
370	375	380	385	
GTC CTA GCA CTG CCA ATC CCC AAG AAG ATG AAA GAG TTC CTC ACA TAC Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr Tyr				1318
390	395	400		

- 138 -

AGG ACT TTC TAGCA GTGCC GGCTCCCCA CCTCCTGCAG CAGCAGCAGT Arg Thr Phe 405	1367
ACAAGGGACT GGCTAGGATG CACTCAGGCA GCTCACACTG GACCAGTGTG GACCTTCCTT 1427	
CCTCCCATGG CATGTGCAAG TAGGTCTGCG TGACCCCCT TCTGTGGTGC CGGCCTTACC 1487	
TCGTCTTCAT CCGTGGTGAG CAGCCTTCGT CAGTCTAGTT GTGTTGAAGC CAAGTGCAGT 1547	
TGTGGATGTT GCTGGGGTAA TAAAGGCAAG CGGGCTCCAG AGCCTCTCTG GTGGCGGCCA 1607	
AGCCACACTC CCTTAACCTG GAAGTACCTG CCACGTAGGG CATTCTGCT GCCTATTTCC 1667	
AGCCAGCGGC TGCATGGTT GAAGTTCCCTC CGTTGTGGTC AGAAGAACCTC TGGTGTGG 1727	
TTCCCTGCTC AGCTGCGCGT GGACTGGGCT GAGCTCCTCA CCATACACTA GTGCCGGCTT 1787	
TTGTTTCCTG TAAACAGTGG TTGCATGTGT AGAGAAGTAA CAAGCGAGTA TTCAGATCAT 1847	
ACGAGGAGGC GTTCCTCGGT GCATGACGGT CAGATGGCCA TTTATCAGCA TATTTATTTG 1907	
TATTTCTCA GCACATAGTA AGGTACAAC GTGTTTCTC AATTGTCTCG AAAAAACAGA 1967	
GTCCTTAAGT GGCCCAGTTG TGGAGCCAAG TCTAAGTCGT GTGGAGTCAG TGCTGACATC 2027	
ACTGGCTTGT GCTGTCTGTC ACATGTGTPT GTCTCTGCTG CTTGACCTCA TGGGATGTAC 2087	
CCTCCTGGTC AACTGCCAA AACAGACAGC CCCTTCCAAG CACCGTTCTT TGACAGCGGT 2147	
ACCAGCTACC TATTCAAGAC GCCTCACACA AAATCTGCCT TAGAAAAGTTA ATATATTTA 2207	
AATTATTTTA AAAGAAAACCTC AACATCTTAT TCTTTGGCCT TTCTTAATTG ATGCTTTATG 2267	
GAGGCAGTGT TAACATTGTA CAGTGTATGC ATAGAGGAGT CTCCTCTATT TGAAGAACAA 2327	
TCCAAAATGA GGCTTTCAATT GAAGGGAAAAA AAAAAAAAAA AA 2369	

(2) INFORMATION FOR SEQ ID NO:21:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 404 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Glu	Ala	Gly	Glu	Glu	Pro	Leu	Leu	Leu	Ala	Glu	Leu	Lys	Pro	Gly
1															15
Arg Pro His Gln Phe Asp Trp Lys Ser Ser Cys Glu Thr Trp Ser Val															
20 25 30															
Ala Phe Ser Pro Asp Gly Ser Trp Phe Ala Trp Ser Gln Gly His Cys															
35 40 45															
Val Val Lys Leu Val Pro Trp Pro Leu Glu Glu Gln Phe Ile Pro Lys															
50 55 60															
Gly Phe Glu Ala Lys Ser Arg Ser Ser Lys Asn Asp Pro Lys Gly Arg															
65 70 75 80															
Gly Ser Leu Lys Glu Lys Thr Leu Asp Cys Gly Gln Ile Val Trp Gly															
85 90 95															
Leu Ala Phe Ser Pro Trp Pro Ser Pro Pro Ser Arg Lys Leu Trp Ala															
100 105 110															
Arg His His Pro Gln Ala Pro Asp Val Ser Cys Leu Ile Leu Ala Thr															
115 120 125															
Gly Leu Asn Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu															
130 135 140															
Leu Leu Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser															
145 150 155 160															
Phe Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys															
165 170 175															
Thr Leu Arg Ile Trp Asp Leu Asn Lys His Gly Lys Gln Ile Gln Val															
180 185 190															
Leu Ser Gly His Leu Gln Trp Val Tyr Cys Cys Ser Ile Ser Pro Asp															
195 200 205															
Cys Ser Met Leu Cys Ser Ala Ala Gly Glu Lys Ser Val Phe Leu Trp															
210 215 220															

- 140 -

Ser Met Arg Ser Tyr Thr Leu Ile Arg Lys Leu Glu Gly His Gln Ser
 225 230 235 240

Ser Val Val Ser Cys Asp Phe Ser Pro Asp Ser Ala Leu Leu Val Thr
 245 250 255

Ala Ser Tyr Asp Thr Ser Val Ile Met Trp Asp Pro Tyr Thr Gly Ala
 260 265 270

Arg Leu Arg Ser Leu His His Thr Gln Leu Glu Pro Thr Met Asp Asp
 275 280 285

Ser Asp Val His Met Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu
 290 295 300

Gly Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp
 305 310 315 320

Ala Leu Glu Leu Lys Ala Pro Val Ala Phe Ala Pro Met Thr Asn Gly
 325 330 335

Leu Cys Cys Thr Phe Phe Pro His Gly Gly Ile Ile Ala Thr Gly Thr
 340 345 350

Arg Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser
 355 360 365

Leu Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr
 370 375 380

Gln Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr
 385 390 395 400

Tyr Arg Thr Phe

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GACACTGCAT CGTCAAACTG ATCCCCGGC CGTTGGAGGA GCAGTTCATC CCTAAAGGGT	60
TTGAAGCCAA AAGCCGAAGT AGCAAAAATG AGACGAAAGG GCGGGGCAGC CCAAAAGAGA	120
AGACGCTGGA CTGTGGTCAG ATTGTCTGGG GGCTGGCCTT CAGCCTGTGC TTTCCCCACC	180
CAGCAGGAAG CTCTGGCAC GCCACCACCC CCAAGTGCCC GATGTCTCTT GCCTGGTTCT	240
TGCTACGGGA CTCAACGATG GGCAGATCAA GATCTGGAG GTGCAGACAG GGCTCCTGCT	300
TTTGAATCTT TCCGGCCACC AAGATGTCGT GAGAGATCTG AGCTTCACAC CCAGTGGCAG	360
TTTGATTTCG GTCTCCCGGT CACGGGATAA GACTCTCGC ATCTGGGACC TGAATAAACAA	420
CGGTAAACAG ATTCAAGTGT TATCGGGCCA CCTGCAGTGG GTTTACTGCT GTTCCATCTC	480
CCCAGACTGC ACCATGCTGT GCTCTGCAGC TCGAGAGAAG TCGGTCTTTC TATGGAGCAT	540
GAGGTCTTAC ACGTTAATTG GGAAGCTAGA GGGCCATCAA AGCAGTGTG TCTCTTGTGA	600
CTTCTCCCC GACTCTGCCG TGCTTGTAC GGCTTCTTAC GATACCAATG TGATTATGTG	660
GGACCTCTAC ACCGGCGAAA GGCTGAGGTC ACTCCACCCAC ACCCAGGTTG ACCCCGCCAT	720
GGATGACAGT GACGTCCACA TTAGCTCACT GAGATCTGTG TGCTTCTCTC CAGAAGGCTT	780
GTACCTTGCC ACGGTGGCAG ATGACAGACT CCTCAGGATC TGGGCCCTGG AACTGAAAAC	840
TCCCATTGCA TTTGCTCCTA TGACCAATGG CCTTTGCTGG CACATTTTT CCACATGGTG	900
GAGTCATTGC CACAGGGACA AGAGATGGCC ACGTCCAGTT CTGGACAGCT CCTAGGGTCC	960
TGTCTCACT GAAGCACTTA TGCCGGAAAG CCCTTCGAAG TTTCTTAACA ACTTACCAAG	1020
TCCTAGCACT GCCAATCCCC AAGAAAATGA AAGAGTTCT CACATACAGG ACTTTTTAAG	1080
CAACACCACA TCTTGTGCTT CTTTGTAGCA GGGTAAATCG TCCTCTCAAA CCCACTTCCT	1140
GGAATAATGG GCCAAACATC TGGTCTTGCA TTGAATAGC ATTTCTTGG GATTGTGAAT	1200

AGAATGTAGC AAAACCAGAT TCCAGTGTAC TAGTCATGGA TTTTTC

1246

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACCATGGTTC CAAGTCCTCT CCCCTGTGGT CAAGTTGCCCGAATGTTGGG CCCAAGTGCC	60
TTTCCTCCT TGGGCCTCCC CTTCTGACCT GCAGGACAGT TTTCCGGAGC CCATTTGGTA	120
TGAGGTATTA ATTAGCCTTA ACTAAATTAC AGGGGACTCA GAGGCCGTGC TCCTGACCGA	180
TCCAGACACT ATTTTTTTTT TTTTTTTTA ACAATGGTGT GCATGTGCAG GAAATGACAA	240
ATTTGTATGT CAGATTATAAC AAGGATGTAT TCTTAAACCG CATGACTATT CAGATGGCTA	300
CTGAGTTATC AGTGGCCATT TATTAGCATC ATATTTATTT GTATTTCTC AACAGATGTT	360
AAGGTACAAC TGTGTTTTC TCGATTATCT AAAAACATA GTACTAAAT TGAAAAAAA	420
AA	422

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2019 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

03052000-102492

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GGCACGAGGC	GGGGTCAGGG	CGGAGGCTGA	GGACCAAGTA	GGCATGGCGG	AGGGCGGGAC	60
CGGCCCGAT	GGACGGGCCG	GCCCCGGACC	CGCAGGT CCT	AATCTGAAGG	AGTGGCTGAG	120
GGAGCAGTTC	TGTGACCATC	CACTGGAGCA	CTGTGACGAT	ACAAGACTCC	ATGATGCAGC	180
CTATGTAGGG	GACCTCCAGA	CCCTCAGGAA	CCTACTGCAA	GAGGAGAGCT	ACCGGAGCCG	240
CATCAATGAG	AAGTCTGTCT	GGTGCTGCGG	CTGGCTTCCC	TGCACACCAC	TGAGGATCGC	300
AGCCACTGCA	GGCCATGGGA	ACTGTGTGGA	CTTCCTCATA	CGCAAAGGGG	CCGAGGTGGA	360
CCTGGTGGAT	GTCAAGGGC	AGACTGCCCT	GTATGTGGCT	GTAGTGAACG	GGCACTTGGA	420
GAGCACTGAG	ATCCTTTTGG	AAGCTGGTGC	TGATCCCAAC	GGCAGCCGGC	ACCACCGCAG	480
CACTCCTGTG	TACCATGCCT	YTCGTGTGGG	TAGGGACGAC	ATCCTGAAGG	CTCTTATCAG	540
GTATGGGCA	GATGTTGATG	TCAACCATCA	TCTGAATTCT	GACACCCGGC	CCCCTTTTC	600
ACGGCGGCTA	ACCTCCTTGG	TGGTCTGTCC	TCTATACATC	AGTGCTGCCT	ACCATAACCT	660
TCAGTGCTTC	AGGCTGCTCT	TGCAGGCTGG	GGCAAATCCT	GAATTCAATT	GCAATGGCCC	720
TGTCAACACC	CAGGAGTTCT	ACAGGGATC	CCCTGGGTCT	GTCATGGATG	CTGTCCTGCG	780
CCATGGCTGT	GAAGCAGCCT	TCGTGAGTCT	GTTGGTAGAG	TTTGGAGCCA	ACCTGAACCT	840
GGTGAAGTGG	GAATCCCTGG	GCCCAGAGGC	AAGAGGCAGA	AGAAAGATGG	ATCCTGAGGC	900
CTTGCAGGTC	TTTAAAGAGG	CCAGAAAGTAT	TCCCAGGACC	TTGCTGAGTT	TGTGCCGGGT	960
GGCTGTGAGA	AGAGCTCTTG	GCAAATACCG	ACTGCATCTG	GTTCCCTCGC	TGCCGCTGCC	1020
AGACCCATA	AAGAAGTTTT	TGCTTATGA	GTAGCATTCA	CATGCAGTGC	TGACTGCAAT	1080
GTGGAAGCCG	ATCACCTGCA	GTGAAAATG	ACACAGACTC	TGGCATCCTG	GGAACCATGG	1140
CCTGTGCTGC	CAGCTTGATC	CTTGGCTGTC	AGTGAAGAAA	AAACGGCTGT	GTTCTCTTGG	1200
ACTGTGATTC	TATCTCAGGT	GCTTGGGCCA	TCGAACGCTC	CTTGAGTCAT	TGTCAACTGA	1260

08092000-1623402

- 144 -

GAGGCACATA CAAACTTAAT TTTGTTCCCTC TTCAGTCTCT CTGTTTGGA TTCTTCCTGG	1320
CAATGTGTGC AGCATGGGCT GAGCCTGGTG ATTGCCCTAG TGGGAAAGGC TTTTTCTCC	1380
AGGCTATGCA TCTATTATG TTCCCTACTTT GCAATTATT GTTCTTTAA GGCTTGATAT	1440
CAAAACAGAA AGAGGTTTGT TAAGAAAAGA TATAGGGAGA AAGGAATTCC GGTTCCGTGC	1500
ACTTGCTAGC CTGCTTTCCCT TGCCCTGGTT TGTCTGTCTA TGCTGCCTGG TGCACATCCC	1560
TTCTCTTIGC TGCCACTGTT CTATTTGGG AGTTGTCTTC CGTCTAACAGAT GGCTTCTGGG	1620
GTTCTATCTT ATTGCACAGA GGTCCCAGAA CAGTGTTCAT AGGGCACCAT CTGCTCTGCC	1680
AAGGGTTTTC TGATGTCTTA CCCTGGGGAT CTTCAGACAG TGGTTACCTT TAGGAGACCC	1740
ACCTGGAACt AACCATTAAAG TGACTGCCCA CATTTCAGATC AGGGACCATC TTAATAGTAC	1800
TCACTGCCAG TCCTCACAAAG AGAAGATGAC ACGGGTGCTC TCTTCAGACA CTCCCATAACA	1860
GGAAAGTTGGA AAATGTCTTG GTCACCTGGG TTGTTCCAG GCTACAACTT CTTGGTGTTC	1920
CACTAARACC AGRATATCCT AGTTTTTGG GTTGACTGTT CCCTCCCCAC TTTCCTTGAA	1980
NCCCAATGCC CNTTTGTKTN GGTTGCTTCC CTAAAATTT	2019

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala Arg Gly Gly Val Arg Ala Glu Ala Glu Asp Gln Val Gly Met Ala			
1	5	10	15

- 145 -

Glu	Gly	Gly	Thr	Gly	Pro	Asp	Gly	Arg	Ala	Gly	Pro	Gly	Pro	Ala	Gly
					20				25					30	
Pro	Asn	Leu	Lys	Glu	Trp	Leu	Arg	Glu	Gln	Phe	Cys	Asp	His	Pro	Leu
					35				40				45		
Glu	His	Cys	Asp	Asp	Thr	Arg	Leu	His	Asp	Ala	Ala	Tyr	Val	Gly	Asp
					50				55			60			
Leu	Gln	Thr	Leu	Arg	Asn	Leu	Leu	Gln	Glu	Glu	Ser	Tyr	Arg	Ser	Arg
					65		70			75			80		
Ile	Asn	Glu	Lys	Ser	Val	Trp	Cys	Cys	Gly	Trp	Leu	Pro	Cys	Thr	Pro
					85				90			95			
Leu	Arg	Ile	Ala	Ala	Thr	Ala	Gly	His	Gly	Asn	Cys	Val	Asp	Phe	Leu
					100				105			110			
Ile	Arg	Lys	Gly	Ala	Glu	Val	Asp	Leu	Val	Asp	Val	Lys	Gly	Gln	Thr
					115				120			125			
Ala	Leu	Tyr	Val	Ala	Val	Val	Asn	Gly	His	Leu	Glu	Ser	Thr	Glu	Ile
					130		135			140					
Leu	Leu	Glu	Ala	Gly	Ala	Asp	Pro	Asn	Gly	Ser	Arg	His	His	Arg	Ser
					145		150			155			160		
Thr	Pro	Val	Tyr	His	Ala	Xaa	Arg	Val	Gly	Arg	Asp	Asp	Ile	Leu	Lys
—	—	—	—	—	—	165			170			175			
Ala	Leu	Ile	Arg	Tyr	Gly	Ala	Asp	Val	Asp	Val	Asn	His	His	Leu	Asn
						180			185			190			
Ser	Asp	Thr	Arg	Pro	Pro	Phe	Ser	Arg	Arg	Leu	Thr	Ser	Leu	Val	Val
					195				200			205			
Cys	Pro	Leu	Tyr	Ile	Ser	Ala	Ala	Tyr	His	Asn	Leu	Gln	Cys	Phe	Arg
					210				215			220			
Leu	Leu	Leu	Gln	Ala	Gly	Ala	Asn	Pro	Asp	Phe	Asn	Cys	Asn	Gly	Pro
					225		230			235			240		
Val	Asn	Thr	Gln	Glu	Phe	Tyr	Arg	Gly	Ser	Pro	Gly	Cys	Val	Met	Asp
					245				250			255			

09062560-400400

- 146 -

Ala Val Leu Arg His Gly Cys Glu Ala Ala Phe Val Ser Leu Leu Val
 260 265 270

Glu Phe Gly Ala Asn Leu Asn Leu Val Lys Trp Glu Ser Leu Gly Pro
 275 280 285

Glu Ala Arg Gly Arg Arg Lys Met Asp Pro Glu Ala Leu Gln Val Phe
 290 295 300

Lys Glu Ala Arg Ser Ile Pro Arg Thr Leu Leu Ser Leu Cys Arg Val
 305 310 315 320

Ala Val Arg Arg Ala Leu Gly Lys Tyr Arg Leu His Leu Val Pro Ser
 325 330 335

Leu Pro Leu Pro Asp Pro Ile Lys Lys Phe Leu Leu Tyr Glu
 340 345 350

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

—

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCATCCATGG CGGAGGGCGG CAGCACGACG GGCGGGCAGG GCCGGGCTCC GCAGGTCGTA	60
ATCTGAAGGA GTGGCTGAGG GAGCAATTGT GTGATCATCC GCTGGAGCAC TGTGAGGACA	120
CGAGGCTCCA TGATGCCAGCT TACGTCGGGG ACCTCCAGAC CCTCAGGAGC CTATTGCAAG	180
AGGAGAGCTA CCGGAGCCGC ATCAACGAGA AGTCTGTCTG GTGCTGTGGC TGGCTCCCCT	240
GCACACCGTT GCGAATCGCG GCCACTGCAG GCCATGGGAG CTGTGTGGAC TTCCCTCATCC	300
GGAAAGGGGGC CGAGGCTGGAT CTGGTGGACG TAAAAGGACA GACGGCCCTG TATGTGGCTG	360

DRAFT - 10/24/97

- 147 -

TGGTGAACGG GCACCTAGAG AGTACCCAGA TCCTTCTCGA AGCTGGCGCG GACCCCAAC	419
--	-----

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 595 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAGGAACAAG AAAAGTGGAC CCTGAGGCCT TGCAGGTCTT TAAAGAGGCC AGAAGTGTTC	60
CCAGAACCTT GCTGTGTCTG TGCCGTGTGG CTGTGAGAAG AGCTCTTGGC AAAACCGGCT	120
TCATCTGATT CCTTCGCTGC CTCTGCCAGA CCCCATAAAG AAGTTTCTAC TCCATGAGTA	180
GACTCCAAGT GCTGCGGTTG ATTCCAGTGA GGGAGAAAGT GATCTGCAGG GAGGTGGACA	240
CCGAGCCCTG AGTGCTGTGC TGCTGCTGGT CTCCCTGATGG CTGTTGCTGC AGAAGATGTC	300
CTCGTAGACT GTCATTGCTC CTCAGGTGCC TGGGCCGCTG AACAGTCCTT GGGTCATTGT	360
CAGCTGAGAG GCTTATACTA AAGTTATTAT TGTTTTCCC AAGTTCTCTG TTCTGGATTT	420
TCAGTTGCAT ATTAATGTAA CGGGCCATGG GGTATGTACA TGTAGGGGCT GAGGTTGGAG	480
GCCTACTAAT TTCCTGTAGG GAAGACTCCC AGCACTTCTG GAACTGTGCT TCTCTTTATT	540
TTTCTACTTC TCAATTGAT GGTCGATTA AAGCCTTCTA GTATCTCAAT GAAAA	595

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 696 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- 148 -

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4..396

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTG ATG TCC GCA ATT CTG AAG GTT GGA CAC CAC TGC TGG CTG CCT GTG	48
Met Ser Ala Ile Leu Lys Val Gly His His Cys Trp Leu Pro Val	
1 5 10 15	
ACA TCC GCT GTC AAT CCC CAA AGG ATG CTG AGG CCA CCA CCA ACC GCT	96
Thr Ser Ala Val Asn Pro Gln Arg Met Leu Arg Pro Pro Pro Thr Ala	
20 25 30	
GTT TTC AAC TGT GCC GCT TGC TGC TGT CTG TGG GGG CAG ATG CTG ATG	144
Val Phe Asn Cys Ala Ala Cys Cys Leu Trp Gly Gln Met Leu Met	
35 40 45	
AAT ACA TAC CGT GTA GTT CAG CTT CCT GAG GAG GCC AAG GGC TTG GTG	192
Asn Thr Tyr Arg Val Val Gln Leu Pro Glu Glu Ala Lys Gly Leu Val	
50 55 60	
CCA CCA GAG ATT CTA CAG AAG TAC CAT GGA TTC TAC TCT TCC CTC TTT	240
Pro Pro Glu Ile Leu Gln Lys Tyr His Gly Phe Tyr Ser Ser Leu Phe	
-65 70 75	
GCC TTG GTG AGG CAG CCC AGG TCG CTG CAG CAT CTC TGC CGT TGT GCG	288
Ala Leu Val Arg Gln Pro Arg Ser Leu Gln His Leu Cys Arg Cys Ala	
80 85 90 95	
CTC CGC AGT CAC CTG GAG GGC TGT CTG CCC CAT GCA CTA CCG CGC CTT	336
Leu Arg Ser His Leu Glu Gly Cys Leu Pro His Ala Leu Pro Arg Leu	
100 105 110	
CCC CTG CCA CCG CGC ATG CTC CGC TTT CTG CAG CTG GAC TTT GAG GAT	384
Pro Leu Pro Pro Arg Met Leu Arg Phe Leu Gln Leu Asp Phe Glu Asp	
115 120 125	
CTG CTC TAC TAGGCTTGCT GCCCTGTGAA CAAAGCAGAC CCCACCCCCA	433
Leu Leu Tyr	
130	

- 149 -

CCCCAAGGGC ATCTCTCAC C AATGAATGAT GCAAGGGCGT CTGTCTCAA GTCAGGAGTG	493
GACGCCCTG A TCCACACTTG AGAGAAGAGG CCAGATCAGC ACCYGGCTGG TAGTGATNGC	553
AGAGGGCACC TGTGCAGATC TGTGTGCGCA CTGGAAATCT CTAGGCTGAA GGCYAGAGCA	613
AATGGTGCA R GTGTTAGTCC TTGGGANGAG AGACAGANGG TGAGAAAGCA AGACAGAGGT	673
GAGAGTGCAC ATGTCAAGTG GTAGATTGCC TTAAAAGAAA GCTAAAAAAA GAAAAAGATT	733
CGGGCGAACT TCTTTAGGGG TAATGCTGCA GCGTGTAAA CTGACTGACC AGCGTCCATA	793
TCTTTGGACC CTTCCCGGGT GAAAAAGCCC CTTCATCCTC CAGCGCTCCC CAAGGGTGCT	853
TAGCAATACC GGGTGCCTTT CTGCCGCAAA GTGAGTTACC AAA	896

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Ser Ala Ile Leu Lys Val Gly His His Cys Trp Leu Pro Val Thr			
1	5	10	15
Ser Ala Val Asn Pro Gln Arg Met Leu Arg Pro Pro Pro Thr Ala Val			
20	25	30	
Phe Asn Cys Ala Ala Cys Cys Cys Leu Trp Gly Gln Met Leu Met Asn			
35	40	45	
Thr Tyr Arg Val Val Gln Leu Pro Glu Glu Ala Lys Gly Leu Val Pro			
50	55	60	
Pro Glu Ile Leu Gln Lys Tyr His Gly Phe Tyr Ser Ser Leu Phe Ala			
65	70	75	80
Leu Val Arg Gln Pro Arg Ser Leu Gln His Leu Cys Arg Cys Ala Leu			
85	90	95	

- 150 -

Arg Ser His Leu Glu Gly Cys Leu Pro His Ala Leu Pro Arg Leu Pro			
	100	105	110
Leu Pro Pro Arg Met Leu Arg Phe Leu Gln Leu Asp Phe Glu Asp Leu			
	115	120	125
Leu Tyr			
	130		

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 436 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTGGGGCGT CATCATGACC TCCTCTAGGG CTCTGCAACA TGACTCCTGT GGTGCAAATC	60
AACATTTGT TCACTGATGA ATCCACAAGG ATCTCTGGGC CTACAACCAG GTCCCTGGTCC	120
ACATGACTGT CGTCTTCGGA GAAGGCACCA CTCGCCCGGC GCAGGTACGG CTGACACCTC	180
CATGGGAGAA GACGTATCCA GGCAGCAGCT GCGCGGCCCT TCAAGAGGGC ACATCCCCTC	240
ATCTAAAGGC ACGGTGTACT GAAGGTAGTC CTGAGACATG AGTCCGATTA CTACAGGCAC	300
GTGTTCTCC AGGTGGAGGC TCAGGTCCCC GGGTGAGCTG GGGCTGCCAGC GGGACTCAGG	360
GCGCGGCTCT GGCTGCAGGT CTCGCAGCTC CCTGGGCTGT AGCTCCCGCA GATCCTTGCG	420
CACACCGTTG ACTGGT	436

(2) INFORMATION FOR SEO ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- 151 -

- (A) LENGTH: 2180 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TTAATAGTAC CTACATAGTA GAAAATTATA ACTCCACTTT AAAACAATGT TTTCTTTCTA	60
TTCAAATCAA TTTAAAAC TT TTTATAAAACA TTAATGTTGC AAGAGAATCC AGTCCATTAA	120
TGAAAATTAG TTGACAATCA AGTTCACCCA AGAAAATGTT GACTAAGCTA AAGAAATCAC	180
AGATAAAAACA TTTTACCAAA AGGATAGGTA ACACACAAAA AAATGCTATC ACAGGAAGCT	240
ATGATCATCT AATATTTCTT TAATAATAAT TCTAGTTCCA TAGGTTTCAT TGTTATGCCA	300
ATTTGTACCC GAGTTTAATT ACAGAAAAGG CAACAATTTC TAAATTGGTG GTATACATT	360
CTTTACAATT TTTTAATGTA AGGCCATTAA TTAAAATAGA CAAACTAGAA GATGAAAACG	420
AAGGCAACAG AAAAATTCAA CTTTCACAA CCAAAAGAAT TAGCACAACC TTAGAAATAA	480
TTTAGAAAAAA AGTGTGTTA AAAGATATGT TGCAGATCTC CGTTCCATTAA CCCAAGATTAA	540
—	
TGTCAATTCA CGATTCTAAA TAAATCTTT TAAAGTAAGA GATTAACAC TCATCTTCAG	600
TGTATATGTA AATTCCGTGG TTTTATCACA CAGGTATGTT TATTCAACAC TGCTTTGGAA	660
ATGGACCATT TAAAAGGACA TGGCAATTTC CATTCTGTTA AGTTTCATTAC AACCTTTACT	720
TAGGGGTTGA TTACCACATG AAATGTGCTT TTAATGCATA AAAATCACAG TGGATTAGCC	780
AGCAAAAGGG ACTGGGCGGG GGGGGCATTG AGGAGAATT GATAATTAC ATTGTGATTA	840
TTCTGCACAT TGATGAAACA TAATTACACAC CTCTAAAACC TCAAGACTTC CCTTTTTAA	900
AGAACCAAAA TAAACCCAAG ACACCTTGCT GACACTTCCC CACCCCTAAA CAAACTGATG	960
ACTCTTTTAC ACATAAAACT GAAATAGTTA TGGCAGCAAA AGATTTGAT GGCAATGAAA	1020

- 152 -

GTTTGTAAC	TGTATTCAA	TCTCTGTC	TTATTCCAA	AGTGCAAGAT	GCAGGGTCT	1080
CAATCTTCA	GTAGTGCTC	TCCTGTAAT	AATCCTTCAT	TTTGTTCGC	AAAGGCAGTT	1140
TCTGAATTAA	GTCTATTCTG	GTATACTGAC	GTATAACAA	ACGACACAGG	TACTGCAACG	1200
AGCGCACCTA	TGAACCCCCG	AACACTGGTT	GGCAAGTTCT	GACGGAAGTG	CAGATTCCAG	1260
GCAGCGAGAC	CTTGAATAAC	AAAAAGCTCC	CATTTCAGA	GTCCCTGATT	GAATGCTCCA	1320
ATTAGATCAA	CTATGGACGT	ATGTCCTTC	ACATCGGCTG	TTCATAAAAG	CTAACCTAC	1380
CATTTGAGTG	CTCAATTCTA	GTGTGAAGTG	TTTACCATG	GGAGCGAAAG	TCACAGCTTA	1440
AAAGGTAACG	GTCGTCAGAA	CTGTCCCGAA	CAAGAARAGA	ACCATCTGGC	ACGTTTGCTA	1500
GCTTCCTTC	TGCCTCCAA	CGTGTGATTG	GTCCCCAGTA	CCATCCTTGC	TTGCAAGTT	1560
TTTCAGCTC	CTCTGTAAGG	CTTGTACAA	CCATGGGACC	ACTACTTG	ACTGAGTCAT	1620
AAACTCTTGC	AACCCCAGGA	GCAGAGTTCG	GATAAAATT	CAAATGACAG	CGCATAACTT	1680
TCAGCCACGT	GGGGCTTTCT	GTCCAGTGAG	TCCACTGAAA	GTTCCCTTT	GGGATTTGGA	1740
TTATT CCTGC	ATTGGAGTAA	CCAATGGTGA	AGATTGGAGG	GACATCCATC	GTGAACCCGC	1800
TCTCCGGGGT	TCTGCAACAT	GAATCCC GTG	GTGCCAATCA	ACAAGCCATT	CACCGGACTG	1860
ATCCAGGAAG	ATCTCTGGGG	CGACAACTAG	GTCCTGGTCT	ACCTGACTCT	CATCCTCGGG	1920
GAAAGCGCGC	CCTCCCAC	T GAGGAGAAC	CGCAGAGACT	TCCATGGGAG	AAGAGCTGTC	1980
CAGACAATAG	CTCCGTGATC	CTTCAAAGG	ATACATCCCC	TCATCTAAAG	GCACAGTATA	2040
CTGAATGTAG	TCCTGAGGCA	TAAGTCAAAT	AACGACAGGC	ACATGTTCAT	CCAGGTGAAG	2100
ATGCAGGTCT	CCATTATGAG	AAGCCGAGCT	CTTCAGTGAA	TTGGCTTGCT	CCTGGCACGT	2160
GGTCTCAGAC	TGGAGGTCGT					2180

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2649 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GGCACGAGGC TGTGTCCAGC ACACAGAGAG GGCCCAGCCA TCTGCTTGG TTCAGAGGCC	60
TGTGTCTGTC TGTCACTTAG ACTCTTCCTC CCGGCTCGCA GCTCACCCCTC CATCCTCCTT	120
ACTGGCTCCA GCATGACTCG CTTCTCTTAT GCAGAGTACT TTGCTCTGTT TCACTCTGGC	180
TCTGCACCTT CCAGGTCCCC TTCGTCTCCC GAGAACCCAC CGGCCCGCGC ACCCCTGGGT	240
CTGTTCCAAG GGGTCATGCA GAAGTATAGC ACCAACCTGT TCAAGACCTC CCAGATGGCG	300
GCTATGGACC CCGTGCTGAA GGCCATCAAG GAAGGGGATG AAGAGGCCTT GAAGATCATG	360
ATCCAGGATG GGAAGAATCT TGCAGAGCCC AACAAAGGAGG GCTGGCTGCC GCTCCACGAG	420
GCTGCCTACT ATGGCCAGCT GGGCTGCCTG AAAGTCCTGC AGCAAGCCTA CCCAGGGACC	480
ATTGACCAAC GCACACTGCA GGAAGAGACA GCATTATACC TGGCCACATG CAGAGAACAC	540
CTGGATTGCC TCCTGTCGCT GCTCCAGGCG GGGCAGAGC CTGACATCTC TAACAAATCC	600
AGGGAGACTC CACTTACAA AGCCTGTGAG CGCAAGAACG CGGAGGCGGT GAGGATATTG	660
GTGCGATACA ACGCAGACGC CAACCACCGC TGTAACAGGG GCTGGACCGC ACTGCACGAG	720
TCTGTCTCCC GCAATGACCT GGAGGTCATG GAGATCCTAG TGAGTGGCGG GGCAAGGTG	780
GAGGCCAAGA ATGTCTACAG CATCACCCCT TTGTTGTGG CTGCCAGAG TGGGCAGCTG	840
GAGGCCCTGA GGTTCTGGC CAAGCATGGT GCAGACATCA ACACGCAGGC CAGTGACAGT	900
GCATCAGCCC TCTACGAGGC CAGCAAGAAT GAGCATGAAG ACGTGGTAGA GTTTCTTCTC	960
TCTCAGGGCG CCGATGCTAA CAAAGCCAAC AAGGACGGCC TGCTCCCCCT GCATGTTGCC	1020

TCCAAGAAGG GCAACTATAG AATAGTGCAG ATGCTGCTGC CTGTGACCGAG CCGCACGCC	1080
GTGCGCCGTA CGGGCATCAG CCCGCTGCAC CTAGCGGCCG AGCGCAACCA CGACGCCGTG	1140
CTGGAGGCGC TGCTGGCCGC GCGCTTCGAC GTGAACGCAC CTCTGGCTCC CGAGCGGCC	1200
CGCCTCTACG AGGACCGCCG CAGTTCTGCG CTCTACTTCG CTGTGGTCAA CAACAATGTG	1260
TACGCCACCG AGCTGTTGCT GCTGGGGGGC GCGGACCCC ACCCGCATGT CATCAGCCCT	1320
CTGCTCGTGG CCATCCGCCA CGGCTGCCTG CGCACCATGC AGCTGCTGTT GGACCATGGC	1380
GCCAACATCG ACGCCTACAT CGCCACTCAC CCCACCGCCT TTCCAGCCAC CATCATGTTT	1440
GCCATGAAGT GCCTGTCGTT ACTCAAGTTC CTTATGGACC TCGGCTGCAG TGGCGAGCCC	1500
TGCTTCTCCT GCCTGTACGG CAACGGGGCG CACCACCCGC CCCGCGACCT GGCGCTTCC	1560
ACGACGCACC CGTGGACGAC AAGGCACCTA GCGTGGTCA GTTCTGTGAG TTCCTGTCGG	1620
CCCCGGAAAGT GAGCCGCTGG GCGGGACCCC TCATCGATGT CCTCCTGGAC TATGTGGGCA	1680
ACGTGCAGCT GTGCTCCCGG CTGAAGGAGC ACATCGACAG CTTTGAGGAC TGGGCTGTCA	1740
TCAAGGAGAA GGCAGAACCT CCGAGACCTC TGGCTCACCT CTGCCGGCTG CGGGTTCGGA	1800
AGGCCATAGG AAAATACCGG ATAAAACCTC TGGACACACT GCCGCTTCCC GGCAGGCTAA	1860
TCAGATACTT GAAATATGAG AATACACAGT AACCAAGCCTG GAGAGGAGAT GTGGCCTTCA	1920
GACTGTTTCC GGGACGCCCG AGGTGGCCTG CATCCAGGAC CCCCTGGGGT CAGAACAGGT	1980
GTGACCTTGC TGGTTCTTTC CTGGAGCTTC ACCCAAAGTG AGAACCTGAT GTGGGGAGTG	2040
GACGTGGAAC CTCTGCTTTC ACACTGTCAG CGGATCCAG ACCCGCTCTG CTTCTGGCCA	2100
TAGCCAGAGA CCTTCAACCT GGGGCCAGGG GAGAGCTGGT CTGGGCAAGG TGGCCCAGGC	2160
AGGAATCCTG GCCTTAAGCT GGAGAACCTG TAGGAATCCC TCACTGGACC CTCAGCTTTC	2220
AGGCTGCGAG GGAGACGCCG AGCCCAAGTA TTTTATTTCG GTGACACAAAT AACGTTGTAT	2280
CAGAAAAAAA AAAAAACATG GGCGCAGCTT ATTCCCTTAGT AGGGTATTAA CTTGCATGCC	2340
CGCTTAAAGC TACTGGAAAC ATGCGTTCCA CTATGCTTGA GAATCCCCCTT GCACTGGTAA	2400

- 155 -

ACGAGAGCCG ACGTGCTTCA AGGTTGGATT TTTGGTTGCC CCTTTGGCGT TCCGCGGTT	2460
TGTCCGACGT AATTGACCCC GTGTTTGTC ACTTTCGAGT GTTCCGACTA TTGGGGGGCT	2520
TTTGGTTGTC CCCAAAATTG TGGGTGGTGT GCGGACGCCA CGAGAAGTGG TTCATGGCG	2580
ATAATCATTA CTGGAGAATG TAGAGCGGCG GTTTTACGAA TAAATATTTT TTAAGCCCC	2640
TTCCCAAAA	2649

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 495 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CCTCCTGAGA GTTCGCCGGC CCGGGCCCAA TGGGTTGTC CAAGGGGTCA TGCAGAAATA	60
CAGCAGCAGC TTGTTCAAGA CCTCCCAGCT GGCGCCTGCC GACCCCTTGA TAAAGGCCAT	120
CAAGGATGCG ATGAAGAGGC CTTGAAGACC ATGATCAAGG AAGGAAAGAA TCTCGCAGAG	180
CCCAACAAGG AGGGCTGGCT GCCGCTGCAC GAGGCCGCAT ACTATGCCA GGTGGGCTGC	240
CTGAAAGTCC TGCAGCGAGC GTACCCAGGG ACCATCGACC AGCGCACCCCT GCAGGAGGAA	300
ACAGCCGTTT ACTTGGCAAC GTGCAGGGGC CACCTGGACT GTCTCCTGTC ACTGCTCCAA	360
GCAGGGGCAG AGCGGGACAT CTCCAACAAA TCCCGAGAGA ACCGCTCTAC AAAGCCTGTG	420
AGCGCAAGAA CGCGGAAGCC GTGAAGATTG TTGGTGCAGC ACAACGCAGA CACCAACAAAC	480
GCTGCAACCG CGCTG	495

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

DRAFT - DRAFT

- 156 -

- (A) LENGTH: 709 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTGCAGCTCT GCTCGCGGCT GAAGGAACAC ATCGACAGCT TTGAGGACTG GGCGTCATC	60
AAGGAGAAAGG CAGAACCTCC AAGACCTCTG GCTCACCTTT GCCGACTGCG GTTTCGAAAG	120
GCCATTGGGA AATACCGTAT AAAACTCCTA GACACCTTGC CGCTCCCAGG CAGGCTGATT	180
AGATAACCTGA AATACGAGAA CACCCAGTAA CTGGGGCCAC GGGGAGAGAG GAGTAGCCCC	240
TCAGACTCTT CTTACTAAGT CTCAGGACGT CGGTGTTCCC AACTCCAAGG GGACCTGGTG	300
ACAGACGAGG CTGCAGGCTG CCTCCCTCTC AGCCTGGACA GCTACCAGGA TCTCACTGGG	360
TCTCAGGGCC CAGAGCTTG GCCAGAGCAG AGAACAGAAT GTGTCAAGGA AAAGAATCAT	420
TTGTTTACAA ACTGATGAGC AGATCCCAGA CCTTCTCTAC CTTCAGGAAT GGCAAGAAACC	480
TCTATTCTG GGGCCAGGGC AGAGCTTGAG GTGTTCTGGG GAAGGTGGTG CTCAGAGCCT	540
—	
TCCCTGTGCC CCTCCACTTG TTCTGGAAAA CTCACCACTT GACTTCAGAG CTTTCTCTCC	600
AAAGACTAAG ATGAAGACGT GGCCCAAGGT AGGGGGTAGG GGGAGCTGG GTCTTGGAGG	660
GCTTTGTTAA GTATTAATAT AATAAATGTT ACACATGTGA AAAAAAAA	709

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:

 - (A) LENGTH: 848 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TTG GAG AAG TGT GGT TGG TAT TGG GGG CCA ATG AAT TGG GAA GAT GCA	48
Leu Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala	
1 5 10 15	
GAG ATG AAG CTG AAA GGG AAA CCA GAT GGT TCT TTC CTG GTA CGA GAC	96
Glu Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp	
20 25 30	
AGT TCT GAT CCT CGT TAC ATC CTG AGC CTC AGT TTC CGA TCA CAG GGT	144
Ser Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly	
35 40 45	
ATC ACC CAC CAC ACT AGA ATG GAG CAC TAC AGA GGA ACC TTC AGC CTG	192
Ile Thr His His Thr Arg Met Glu His Tyr Arg Gly Thr Phe Ser Leu	
50 55 60	
TGG TGT CAT CCC AAG TTT GAG GAC CGC TGT CAA TCT GTT GTA GAG TTT	240
Trp Cys His Pro Lys Phe Glu Asp Arg Cys Gln Ser Val Val Glu Phe	
65 —— 70 75 80	
ATT AAG AGA GCC ATT ATG CAC TCC AAG AAT GGA AAG TTT CTC TAT TTC	288
Ile Lys Arg Ala Ile Met His Ser Lys Asn Gly Lys Phe Leu Tyr Phe	
85 90 95	
TTA AGA TCC AGG GTT CCA GGA CTG CCA CCA ACT CCT GTC CAG CTG CTC	336
Leu Arg Ser Arg Val Pro Gly Leu Pro Pro Thr Pro Val Gln Leu Leu	
100 105 110	
TAT CCA GTG TCC CGA TTC AGC AAT GTC AAA TCC CTC CAG CAC CTT TGC	384
Tyr Pro Val Ser Arg Phe Ser Asn Val Lys Ser Leu Gln His Leu Cys	
115 120 125	
AGA TTC CGG ATA CGA CAG CTC GTC AGG ATA GAT CAC ATC CCA GAT CTC	432
Arg Phe Arg Ile Arg Gln Leu Val Arg Ile Asp His Ile Pro Asp Leu	
130 135 140	

PROGRESS REPORT

- 158 -

CCA CTG CCT AAA CCT CTG ATC TCT TAT ATC CGA AAG TTC TAC TAC TAT Pro Leu Pro Lys Pro Leu Ile Ser Tyr Ile Arg Lys Phe Tyr Tyr Tyr	480
145 150 155 160	
 GAT CCT CAG GAA GAG GTA TAC CTG TCT CTA AAG GAA GCG CAG CGT CAG Asp Pro Gln Glu Glu Val Tyr Leu Ser Leu Lys Glu Ala Gln Arg Gln	528
165 170 175	
 TTT CCA AAC AGA AGC AAG AGG TGG AAC CCT CCA CGT AGC GAG GGG CTC Phe Pro Asn Arg Ser Lys Arg Trp Asn Pro Pro Arg Ser Glu Gly Leu	576
180 185 190	
 CCT GCT GGT CAC CAC CAA GGG CAT TTG GTT GCC AAG CTC CAG CTT TGAAGAACCA 631	
Pro Ala Gly His His Gln Gly His Leu Val Ala Lys Leu Gln Leu	
195 200 205	
 AATTAAGCTA CCATGAAAAG AAGAGGAAAAA GTGAGGGAAC AGGAAGGTTG GGATTCTCTG	691
TGCAGAGACT TTGGTTCCCC ACGCAAGCCC TGCGGCTTGG AAGAAGCACA TGACCGTACT	751
CTGCCTGGGG CTCCACCTCA CACCCACCCC TGCGCATCTT AGGACTGGAG GGGCTCCTTG	811
GAAAATGGA AGAAGTCTCA ACACTGTTTC TTTTCA	848

(2) INFORMATION FOR SEQ ID NO:36:

---(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Leu Glu Lys Cys Gly Trp Tyr Trp Gly Pro Met Asn Trp Glu Asp Ala 1 5 10 15	
 Glu Met Lys Leu Lys Gly Lys Pro Asp Gly Ser Phe Leu Val Arg Asp 20 25 30	
 Ser Ser Asp Pro Arg Tyr Ile Leu Ser Leu Ser Phe Arg Ser Gln Gly 35 40 45	

- 159 -

Ile	Thr	His	His	Thr	Arg	Met	Glu	His	Tyr	Arg	Gly	Thr	Phe	Ser	Leu
50															60
Trp	Cys	His	Pro	Lys	Phe	Glu	Asp	Arg	Cys	Gln	Ser	Val	Val	Glu	Phe
65															80
Ile	Lys	Arg	Ala	Ile	Met	His	Ser	Lys	Asn	Gly	Lys	Phe	Leu	Tyr	Phe
															95
Leu	Arg	Ser	Arg	Val	Pro	Gly	Leu	Pro	Pro	Thr	Pro	Val	Gln	Leu	Leu
															110
Tyr	Pro	Val	Ser	Arg	Phe	Ser	Asn	Val	Lys	Ser	Leu	Gln	His	Leu	Cys
															125
Arg	Phe	Arg	Ile	Arg	Gln	Leu	Val	Arg	Ile	Asp	His	Ile	Pro	Asp	Leu
															130
Pro	Leu	Pro	Lys	Pro	Leu	Ile	Ser	Tyr	Ile	Arg	Lys	Phe	Tyr	Tyr	Tyr
															145
Asp	Pro	Gln	Glu	Glu	Val	Tyr	Leu	Ser	Leu	Lys	Glu	Ala	Gln	Arg	Gln
															165
Phe	Pro	Asn	Arg	Ser	Lys	Arg	Trp	Asn	Pro	Pro	Arg	Ser	Glu	Gly	Leu
															180
Pro	Ala	Gly	His	His	Gln	Gly	His	Leu	Val	Ala	Lys	Leu	Gln	Leu	
															195
															200
															205

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

- 160 -

GTTCCAAGCC TAACCCATCT TTGTCGTTG GAAATTGGGG CCAGTCTAAA AGCAGAGCAC	60
CTTCACTCTG ACATTTCAT CCATCAGTTG CCACTTCCC AAGTCTGCA GAACTATTTG	120
CTCTATGAAG AGGTTTAAG AATGAATGAG ATTCTAGAAC CAGCAGCTAA TCAGGATGGA	180
GAAACCAGCA AGGCCACCTG ACACAGGTCC TTTAATTCTG TTTACTCACA AAAGACGGCT	240
TGTGTGACTG TTTGGATTTG GTGATCAAAT GTCCATGTT ACAGTTGCTT TTCCAGTTT	300
GTCCTTTCC CAATATTGTG AACCTTATCC ATCTTGCCTT ACTCAGTTT ATTCTAGTG	360
CACTTGTG TGTTACCTGA CCATTTCTA CTTTATTCTG CTAATAAACT	420
GTAATTCTGA AAAAAAAA AAAAAAAA AAAAAAAA AAAA	464

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 747 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGGGATCGAA AGCGGGGGCT TCTGGGACGC AGCTCTGGAG ACGCGGCCTC GGACCAGCCA	60
TTTCGGTGTGAAAGTGGCAG CACGGCAGAC TGGTCAAACA AATGGATTTT ACAGAGGCTT	120
ACGCGGACAC GTGCTCTACA GTTGGACTTG CTGCCAGGGAGGCAATGTT AAAGTCTTAA	180
GGAAACTGCT CAAAAAGGGC CGAAGTGTG ATGTTGCTGA TAACAGGGAGGATGCCAA	240
TTCATGAAGC AGCTTATCAC AACTCTGTAG AATGTTGCA AATGTTAATT AATGCAGATT	300
CATCTGAAAA CTACATTAAG ATGAAGACCT TTGAAGGTTT CTGTGCTTTG CATCTCGCTG	360
CAAGTCAAGG ACATTGGAAA ATCGTACAGA TTCTTTAGA AGCTGGGGCA GATCCTAATG	420
CAACTACTTT AGAAGAAACG ACACCAATTGT TTTAGCTGT TGAAAATGGA CAGATAGATG	480

- 161 -

TGTTAAGGCT GTTGCTTCAA CACGGAGCAA ATGTTAACGG ATCCCATTCT ATGTGTGGAT	540
GGAACCTCCTT GCACCAGGCT TCTTTTCAGG AAAATGCTGA GATCATAAAA TTGCTTCTTA	600
GAAAAGGAGC AAACAAGGAA TGCCAGGATG ACTTTGGAAT CACACCTTA TTTGTGGCTG	660
CTCACTATGG CCAAGCTAGA AAGCTTGAA GCATACTTAT TTCATCCGGG TGCAAATGTC	720
AATTGTCAAG CCTTGGACAA AGCTACC	747

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1018 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CACAAATGGG ACCATACAAA AATCTTGGAC TTGTTAATAA CCACTTACTA ACCGGGACCT	60
GTGACACTGG GCTAAACAAA GTAAAGTCCCT GTTTACTCAG CAGTGTGTTGG GGGACATGAA	120
GGATT EE CTA GAAATATTAC TCCGGAATGG TCTACAGCCC AGACGCCAG CGGTGCCTTG	180
TTTTTGGATT CAGTTCTCCT GTGTGCATGG CTTTCCAAAA GGAGGTGGAG CTGTAGTTCT	240
TTGGAATTGT GAACATTCTT TTGAAATATG GAGCCCAGAT AAATGAACCTT CATTTGGCAT	300
ACTGCCTGAA GTACGAGAAG TTTTCGATAT TTGCTACTT TTTGAGGAAA GGTTGCTCAT	360
TGGGACCATG GAACCATATA TATGAATTG TAAATCATGC AATTAAAGCA CAAGCAAAAT	420
ATAAGGAGTG GTTGCACAT CTTCTGGTTG CTGGATTTGA CCCACTGATT CTACTGTGCA	480
ATTCTTGGAT TGACTCAGTC AGCATTGACA CCCTTATCTT CACTTTGGAG TTTACTAATT	540
GGAAGACACT TGCACCAGCT GTTGAAAGGA TGCTCTCTGC TCGTGCCTCA AACGCTTGGA	600
TTCTACAGCA ACATATTGCC CACTGTTCCA TCCCTGACCC ATCTTTGTGG TTTGGAAATT	660

CGGTCCAGTC TAAAATCAGA ACGTCTACGG TCTGACAGTT ATATTAGTCA GCTGCCACTT	720
CCAGAACGCC TACATAATTA TTTGCTCTAT GAAGACGTTG TGAGGATGTA TGAAGTTCCA	780
GAACCTGGCAG CTATTCAAGA TGGATAAAATC AGTGAAACTA CTTAACACAG CTAATTTTT	840
TCTCTGAAAA ATCATCGAGA CAAAAGAGCC ACAGACTACA AGTTTTATG ATTTTATAGT	900
CAAAAGATGA TTATTGATTG TCAGATAGGT TAGGTTTTGG GGGGCCAGTA GTTCAGTGAG	960
AATGTTTATG TTTACAACTA GCCTTCCCAG TAAAAAAAAA AAAAAAAAAA AAAAAAAA	1018

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1897 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CGGGGGGCTG GGACCTGGGG CGTAACCGTC TCTACCACGA CGGCAAGAAC CAGCCAAGTA	60
AAACATACCC AGCCTTTCTG GAGCCGGACG AGACATTCAAT TGTCCTGAC TCCTTTTCG	120
TGGCCCTGGA CATGRATGAT CGGACCTTAA GTTTCATCGT GGATGGACAG TACATGGAG	180
TGGCTTTCCG GGGACTCAAG GGTAAAAAGC TGTATCCTGT ACTGACTGCC GTCTGGGCC	240
ACTGTGAGAT CCGCATGCGC TACTTGAACG GACTTGATCC TGAGCCCCCTG CCACTCATGG	300
ACCTGTGCCG GCGTTCGGTG CGCCTAGCGC TGGAAAAGA CGCCCTGGGT GCCATCCCCG	360
CTCTGCCGCT ACCTGCCTCC CTCAAAGCCT ACCTCCTCTA CCAGTGATCC ACATCCCAGG	420
ACCGCCATAC GACAGCCATC TGGTGCCAAR TCACTGAGCC CGTTGGGTC CGCCGACCCC	480
TGCGCCTGGG ATGGAAGCCC ACCTCAGCCA TGGGCAGACG TGCCCCCTCA TCCTACCGGC	540

TGCCTCTGCT	GGGGGAACCT	ATGCCAACGG	ACTTCTCCCT	TCCCAACACT	GGCTGAAGCA	600
GCAGCACCCA	GGCCCTTCCC	TGAACCAGAT	GCAGAGAATA	AACTATGAAA	ACCTCTCTCA	660
GGCGCCTTCT	GCTCTCAGGT	GGAGTGGGCT	GCCCCCCACT	CTCTGCAGAG	AGAGGCTACA	720
CCCACCTGGG	GGGTCTGGG	AGGTAAGACT	AGTAGGAGGT	GCCAGGGCTG	ARTCCAAAAG	780
CAGGAATGGC	CAGGAMCAGG	CCATACAGAT	GAAGCTCAGG	ATGTCACATA	CCATGGACAM	840
TGAGACAGAA	CCCCAGGTTG	GAMTTCCCTT	GGGCCAACGA	GTGCCAGCTT	TAATGTCAGC	900
TGCMGGTGCT	CTGTGGCCTG	TATTTATTCT	TTAACAGTA	GCAAAGGCCA	TTTATTTATT	960
CCACTTAGAA	AGGAAACCTT	GGTGGGTGGY	TTCCCTCGAT	GTGCTTCCTT	CCACCTCCCT	1020
GGAATGTGTG	TGCCACACCT	GTCCTTGTCC	CAGGCCAGGA	CTGTGGCACA	TGAGCTGGTG	1080
TGCACAGATA	CACGTATGTC	GTCGTGCATG	ACCCCTGACT	AGTTCTAAG	TAGCCCTGCA	1140
CCAAGCACCA	GAGCAGACCC	CAAGAGAGGC	CCGTGCAAGT	CCCCATGTCC	CCAGGTCCCT	1200
GCTTCTGTTG	CCTTGGGACT	CATACACCGG	CACACGTGTT	TCAGCCTCTT	GACTTCCATG	1260
AGCTTCGAAT	TTTGCCCCCG	ATTCTCTGA	TATTTCCCAT	TGGCATCCTC	CAAAGCTCTG	1320
GGCCTGGAGG	GCATTAGGAC	ACATGGAATG	AGTGGGGTCT	CCAGCCCTG	GGAAAGCCAC	1380
TGGCAAGGCA	GGATTAGAAA	GACCAAGAGC	AGGGTGGGGC	GCCATGAAGC	CTGTATGCCT	1440
CTCAGGCTCA	AGACCCCGCC	ACACACCCAC	TCAAGCCTCA	GAAGTGGTGT	GTAGGGCAGC	1500
CCCAGGAGAG	GAATGCCTGT	CCTAGCAGCA	CGTACATGGA	GCACCCACA	TGTGCTCCAG	1560
CCCTCTGGCT	GTTCCTCTTG	CTCTAGAATC	AACTCCCTAC	ATTGGGAATG	TAGCCATTG	1620
GTAGAGGACT	TGCCTAGCCT	GCAGGAAGCT	CACGTTCCAT	CCCCTGCACC	AAGGAGAAC	1680
AAAGCTCAGG	AGGCTGAGGC	AGGAGGATTG	CTGTCAGTGG	TGTACAGAGG	TCATGGCCAT	1740
CCTGGGCTAT	ATTAACACCTT	GTCCTTTAAG	AAAAAGAAAA	CAAATCAACT	TCCATTGAAT	1800
CTGAGTTCTG	CTCATTCTG	CACAGGTACA	ATAGATGACT	TKATTGTTG	AAAAATGKTT	1860
AATATATTTA	CMTATATATA	TATTTGTAAG	AAGCATT			1897

- 164 -

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Gly	Gly	Trp	Asp	Leu	Gly	Arg	Asn	Arg	Leu	Tyr	His	Asp	Gly	Lys	Asn
1				5					10					15	
<hr/>															
Gln	Pro	Ser	Lys	Thr	Tyr	Pro	Ala	Phe	Leu	Glu	Pro	Asp	Glu	Thr	Phe
				20					25				30		
<hr/>															
Ile	Val	Pro	Asp	Ser	Phe	Phe	Val	Ala	Leu	Asp	Met	Xaa	Asp	Gly	Thr
				35				40				45			
<hr/>															
Leu	Ser	Phe	Ile	Val	Asp	Gly	Gln	Tyr	Met	Gly	Val	Ala	Phe	Arg	Gly
				50			55			60					
<hr/>															
Leu	Lys	Gly	Lys	Leu	Tyr	Pro	Val	Val	Ser	Ala	Val	Trp	Gly	His	
	65			70				75				80			
<hr/>															
Cys	Glu	Ile	Arg	Met	Arg	Tyr	Leu	Asn	Gly	Leu	Asp	Pro	Glu	Pro	Leu
				85				90				95			
<hr/>															
Pro	Leu	Met	Asp	Leu	Cys	Arg	Arg	Ser	Val	Arg	Leu	Ala	Leu	Gly	Lys
				100				105				110			
<hr/>															
Glu	Arg	Leu	Gly	Ala	Ile	Pro	Ala	Leu	Pro	Leu	Pro	Ala	Ser	Leu	Lys
	115				120							125			
<hr/>															
Ala	Tyr	Leu	Leu	Tyr	Gln										
		130													

(2) INFORMATION FOR SEQ ID NO:42:

- 165 -

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAGGGTAAAA AACTGTATCC TGTAGTGAGT GCCGTCTGGG GCCACTGTAG ATCCGAATGC	60
GCTACTTGAA CGGACTCGAT CCCGAGACTG CCGCTCATGG ATTTGTGCCG TCGCTCGGTG	120
CGCCTGGCCC TGGGGAGGGA GCGCCTGGGG GAGAACACACA CCTGCCGCTG CCGGCTTCCC	180
TCAAGGCCTA CCTCCTCTAC CAGTGACGTT CGCCATCATA CCGCCAGCGC GACAGCCACC	240
TGGTGCCAAC TCACTGAGCC GCCTG	265

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2438 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AAGTGGCGGC GGTCCCTGGA GAGCAGGCGG AGGCAGCGGC AAGTCTGACT CTGGGCTGAC	60
CGTGGAGCCG GGGCGGGGGC TGACAGCCAG GCCTCCGCCT GGCGGGAGCC GCACGAGGAG	120
CGGGACTGGC CGGGCCTCTC TTCCGGCTT GAGCGAGCGC CGGGTGATGG CGGTGGTGAT	180
GGCGGGCAGGC GCTCGGACAG CTCCGCTTGA GCTGAGCTCG GAGAGATCCG TCCAGAAAGT	240

GCCCAGAAGA AACTTCCTCT TAGAAAAGCT GAAAAACACA RTATTTATAA CACTGAAAT	300
TGTAAAGAAT TTGTTTAAAA TGGCTGAAAA CAATAGTAAA AATGTAGATG TACGGCCTAA	360
AACAAGTCGG AGTCGAAGTG CTGACAGGAA GGATGGTTAT GTGTGGAGTG GAAAGAAGTT	420
GTCTTGGTCC AAAAAGAGTG AGAGTTGTTG TGAATCTGAA GCCATAGGTA CTGTTGAGAA	480
TGTTGAAATT CCTCTAAGAA GCCAAGAAAG GCAGCTTAGC TGTCGTCCA TTGAGTTGGA	540
CTTAGATCAT TCCTGTGGGC ATAGATTTT ACCCGATCC CTTAACACAA AACTGCAAGA	600
TGCGGTGGGG CAGTGTTC CAATAAAGAA TTGTAGTGGC CGACACTCTC CAGGGCTTCC	660
ATCTAAAAGA AAGATTCTA TCAGTGAACAT CATGTTAGAT AAGTGCCCTT TCCCACCTCG	720
CTCAGATTTA GCCTTTAGGT GCCATTAT TAAACGACAC ACTGTTCTA TGAGTCCCAA	780
CTCAGATGAA TGGGTGAGTG CAGACCTGTC TGAGAGGAAA CTGAGAGATG CTCAGCTGAA	840
ACGAAGAAAC ACAGAACATG ACATACCCCTG TTTCTCACAT ACCAATGGCC AGCCTGTGT	900
CATAACTGCC AACACTGCTT CGTGTACAGG TGGTCACATA ACTGGTTCTA TGATGAACCTT	960
GGTCACAAAC AACAGCATAG AAGACAGTGA CATGGATTCA GAGGATGAAA TTATAACGCT	1020
GTGCACAAGC TCCAGAAAAA GGAATAAGCC CAGGTGGAA ATGGAAGAGG AGATCCTGCA	1080
GTTGGAGGCA CCTCTTAAGT TCCACACCCA GATCGACTAC GTCCACTGCC TTGTTCCAGA	1140
CCTCCCTTCAG ATCAGTAACA ATCCGTGCTA CTGGGGTGTC ATGGACAAAT ATGCAGCCGA	1200
AGCTCTGCTG GAAGGAAAGC CAGAGGGCAC CTTTTTACTT CGAGATTCA CGCAGGAAGA	1260
TTATTTATTC TCTGTTAGTT TTAGACGCTA CAGTCGTTCT CTTCATGCTA GAATTGAGCA	1320
GTGGAATCAT AACTTTAGCT TTGATGCCCA TGATCCTGT GTCTTCCATT CTCCTGATAT	1380
TACTGGGCTC CTGGAACACT ATAAGGACCC CAGTGCCTGT ATGTTCTTG AGCCGCTCTT	1440
GTCCACTCCC TTAATCCGGA CGTTCCCCCTT TTCCCTGCAAG CATATTTGCA GAACGGTTAT	1500
TTGTAATTGT ACGACTTACG ATGGCATCGA TGCCCTTCCC ATTCCCTCGC CTATGAAATT	1560
GTATCTGAAG GAATACCATT ATAAATCAAAGT TAGGTAACTCAGGATTG ATGTGCCAGA	1620

- 167 -

GCAGCAGTGA TGC GGAGAGG TTAGAATGTC GACCTGCATA CATATTTCA TTTAATATTT	1680
TATTTTCTT ATGCCCTCTT GAATTTTGTA ACAAAGGCAG TTGAATCAA AAAACTGTG	1740
CCCTAAGTTT TAATTCCAGA TCAATTATT TTTTTATGA TACACTGTT ATATATTTT	1800
AAGCAGGTGT TTGGTTTGTT TTTTACCAT AAAATTACA TATGGTCCAG GCATATTTAC	1860
AATTTCAAGG CATTGGATAT ACATTGAAT ATTCTGTATT TTTAAATAA TCTTTGTTC	1920
TTTCCTATGT CTGAAATATT TTGCTAATCT ATGCTATCAG TATTCTGTA TGACCGAATA	1980
GTTACCTATT CTCTTTCAT CTTGAAGATT TTCAGTAAAG AGTGTGTAA TCAATCCATT	2040
ATAATGTAAT TGACTTTGT AATTCGCAA TAGGAGTGT AAACAACAAA ATGATTAAA	2100
ATGAAACTTA ATGTATTTC ATTTAAATA TTAACAAAC CAAGTTGTT TGTTAGTTAT	2160
TCTAGCCAAT AAGAAAAGAG AATGTAGCAT CCTAGAGGTG TATTTGTCT GCAGTTGGC	2220
AGGACCGTCA GTTAGTCAA ATAAACATCC CCTCAGCGTG GAGGCGAATG GAACCTGTGC	2280
TCCTTTCTTA CGGGAAGCTT TGCAAAGCAA AATAGCAGGG TTACAAGCTT GGAGTTGTTA	2340
AGGCAACTAG AGTTTTCTCT ATTAATTAT AGACTGTTGT TGCACCTACT TAGCTCTTTT	2400
TTGGGAACTC TAGTTCCCAG GGGAAAATAC CTCGTGCC	2438

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 542 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ser	Gly	Gly	Gly	Pro	Trp	Arg	Ala	Gly	Gly	Ser	Gly	Lys	Ser	Asp
1				5				10				15		

- 168 -

Ser	Gly	Leu	Thr	Val	Glu	Pro	Gly	Arg	Gly	Leu	Thr	Ala	Arg	Pro	Pro
		20			25			30							
Pro	Gly	Gly	Ser	Arg	Thr	Arg	Ser	Gly	Ser	Gly	Arg	Ala	Ser	Leu	Pro
		35			40			45							
Arg	Leu	Ser	Glu	Arg	Arg	Val	Met	Ala	Val	Val	Met	Ala	Ala	Gly	Ala
		50			55			60							
Arg	Thr	Ala	Pro	Leu	Glu	Leu	Ser	Ser	Glu	Arg	Ser	Val	Gln	Lys	Val
		65			70			75			80				
Pro	Arg	Arg	Asn	Phe	Leu	Leu	Glu	Lys	Leu	Lys	Asn	Thr	Xaa	Phe	Ile
			85				90				95				
Thr	Leu	Glu	Ile	Val	Lys	Asn	Leu	Phe	Lys	Met	Ala	Glu	Asn	Asn	Ser
			100				105			110					
Lys	Asn	Val	Asp	Val	Arg	Pro	Lys	Thr	Ser	Arg	Ser	Arg	Ser	Ala	Asp
			115				120			125					
Arg	Lys	Asp	Gly	Tyr	Val	Trp	Ser	Gly	Lys	Leu	Ser	Trp	Ser	Lys	
			130				135			140					
Lys	Ser	Glu	Ser	Cys	Ser	Glu	Ser	Glu	Ala	Ile	Gly	Thr	Val	Glu	Asn
			145			150			155			160			
Val	Glu	Ile	Pro	Leu	Arg	Ser	Gln	Glu	Arg	Gln	Leu	Ser	Cys	Ser	Ser
					165			170			175				
Ile	Glu	Leu	Asp	Leu	Asp	His	Ser	Cys	Gly	His	Arg	Phe	Leu	Gly	Arg
					180			185			190				
Ser	Leu	Lys	Gln	Lys	Leu	Gln	Asp	Ala	Val	Gly	Gln	Cys	Phe	Pro	Ile
			195				200			205					
Lys	Asn	Cys	Ser	Gly	Arg	His	Ser	Pro	Gly	Leu	Pro	Ser	Lys	Arg	Lys
			210				215			220					
Ile	His	Ile	Ser	Glu	Leu	Met	Leu	Asp	Lys	Cys	Pro	Phe	Pro	Pro	Arg
			225				230			235			240		
Ser	Asp	Leu	Ala	Phe	Arg	Trp	His	Phe	Ile	Lys	Arg	His	Thr	Val	Pro
					245			250			255				

DECODED PROTEIN

A

- 169 -

Met Ser Pro Asn Ser Asp Glu Trp Val Ser Ala Asp Leu Ser Glu Arg			
260	265	270	
Lys Leu Arg Asp Ala Gln Leu Lys Arg Arg Asn Thr Glu Asp Asp Ile			
275	280	285	
Pro Cys Phe Ser His Thr Asn Gly Gln Pro Cys Val Ile Thr Ala Asn			
290	295	300	
Ser Ala Ser Cys Thr Gly Gly His Ile Thr Gly Ser Met Met Asn Leu			
305	310	315	320
Val Thr Asn Asn Ser Ile Glu Asp Ser Asp Met Asp Ser Glu Asp Glu			
325	330	335	
Ile Ile Thr Leu Cys Thr Ser Ser Arg Lys Arg Asn Lys Pro Arg Trp			
340	345	350	
Glu Met Glu Glu Glu Ile Leu Gln Leu Glu Ala Pro Pro Lys Phe His			
355	360	365	
Thr Gln Ile Asp Tyr Val His Cys Leu Val Pro Asp Leu Leu Gln Ile			
370	375	380	
Ser Asn Asn Pro Cys Tyr Trp Gly Val Met Asp Lys Tyr Ala Ala Glu			
385	390	395	400
Ala Leu Leu Glu Gly Lys Pro Glu Gly Thr Phe Leu Leu Arg Asp Ser			
405	410	415	
Ala Gln Glu Asp Tyr Leu Phe Ser Val Ser Phe Arg Arg Tyr Ser Arg			
420	425	430	
Ser Leu His Ala Arg Ile Glu Gln Trp Asn His Asn Phe Ser Phe Asp			
435	440	445	
Ala His Asp Pro Cys Val Phe His Ser Pro Asp Ile Thr Gly Leu Leu			
450	455	460	
Glu His Tyr Lys Asp Pro Ser Ala Cys Met Phe Phe Glu Pro Leu Leu			
465	470	475	480
Ser Thr Pro Leu Ile Arg Thr Phe Pro Phe Ser Leu Gln His Ile Cys			
485	490	495	

- 170 -

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4999 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CCCTCTGGGC AAGCCGCCCC CCCCCCACCC ATCTACCACA CACACACACA CACACACACA	60
CACACATTCA GACCTTGGGG CAAAAACAAA GC AAAATAAC ACAACAAAA ACAC TGCCTG	120
—	
TGGAAAGTCC TTACTTCAGG AAGGTTGGCA GATGAGGAGC AAGGGAACAT TTTATCAGGA	180
CTGCCACAAA GGAGTCTTT TTTTTAATGG TTTTCAGA CAGGGTTCT CTGTATAGCC	240
CTGGCTGTCC TGGAGCTCAC TTTGTAGACC AGGCTGGCCT CGAACTCAGA AATTGCCTG	300
CCTCTGCCTC CTGAGTGCTG GGATTAAGG CGTGCAGCAC CATGTCCAAC TGGCATTTC	360
TCAATTAAGG TTCGTTCCCT TCAGATAACT CTAGGTTCTG GGTCAAGCTG ACACAAGGCT	420
ACACAGCACA GTTGTATGC CACATTCACT TCAGAAGACA CCCAACCTCC CTGGA ACTGG	480
AACTTATGCA CATTGTGAG CTTCCACTTG GGAGTGGAA CCTGA ACTGG GTCCTCTGCA	540
AGAGCAGCCG TGCTCTTAAC TGCTGAGCCA TTTCAGCAGC CTCACATCAG AATTAAAGTTA	600

GAAATTAGCCG	GGTATGAATC	ATACCCCTAG	AATCCTAGCA	TCTGAAACCA	GAGCTAAGAG	660
AAACAGGGAT	TCAAGACCAG	CTCTGGCTA	CAGAGCCCGT	CCTGTCTAG	GATGGGCTAC	720
AAGAGACTAT	TTCAAAGCCA	TCCAAACAAAC	AATAACTACA	ACAACAACAA	GGTTAAAATT	780
AGGCTGGGCA	CAGGGTACAC	ACCTTTAATG	CCAACACTCA	GGAGGCAGAG	GCAGGCTGAT	840
CAGTGTGAGT	TTGAGTTCAA	CGTGGTCTAC	ATAGGGAGTT	CTAGGCCAGC	AGAGGTTACA	900
GTCTCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTCTCACACA	CACACACACA	CACACACACA	960
CACACACACA	CACACACCGT	GGCATTATGG	GATTTTTTG	GGATAAGGTT	TCTCTGTCTA	1020
GCCCTGGCAT	AGATTCACTC	TGTAGACTAG	GCTAGCCTTG	AACTCAGAGA	TCCGCCTGCC	1080
TCTGCCTCCC	AACTGCTGGG	ATTATAGGTG	TTGCACCACC	ACTGCCAGC	CACTTTGGGA	1140
TTTTGAACT	GTTATCAAGA	GGCTTCGAG	GAGGTCAAAC	TTCAACAGCA	ACCTCTCCAT	1200
GATAATGTAG	CTAATGATCA	AACGACACTC	AAAACCTAAC	CCTTAAAGCA	CACATCCACC	1260
AGACAGCGTG	CCCACTCGTA	GTTCCATTAC	TCAGGAGGCT	GAAGCAGGAG	GATGAAGGAC	1320
TAAGGCTTCA	GCAACCTAGG	GAGCCGCAGG	GGACAGTAGT	CTCAATCCCT	ACATTCTCCT	1380
GAACACAGGA	GCAGGAGTTC	AGGAAGGGTG	TCAAGGCCGC	TTACTGATCT	TAGGCCCTCA	1440
CGAATGACTA	GCTCAGGCAG	AGAGAGCAA	GGTCTCCAGT	GGAGAACTCT	ACACACACAC	1500
ACACACACAC	ACACACACAC	ACACACACAC	AGAATCCAAG	GCGATGACGT	CATCAAAGGG	1560
TTAATTCTAG	TCTGGGATGG	GGGGGAGGGT	GGGGCACCCA	GCTGTCAGGT	GGCTTTGGAA	1620
AAATAAAACTG	CTGAAGAGTC	TGACGCCAGG	GAGTCCTGGG	AGGGACAAGA	GGTTACCCAC	1680
TCAAAGAGTG	TGCTCCACAA	AGCATGCGCG	CTTGTCCACG	TCTGGAGTCG	TCACTTATTT	1740
TTTGCCTGGA	TTCTTTGTAG	CCGGTGGGTT	CTCAAGGCAG	TAAGTGGTGT	GGCCGCCGTG	1800
GTCTGGGAGG	TGACGATAGG	GTAAATCGTC	CACAGAGCCC	AGGGGCGGAG	CGCGGGCGGG	1860
CGTCCGCAGC	CCCGCTGGAG	CCGGAAGCAG	TGGCTGGTCA	GGGGCGCTTC	TAGCCTTCCC	1920
TATCTGTACT	TCCACAGAGG	TCTCTGCGAG	CTAGGGGGAC	AGTGAGGTGC	GGGGTAGGGG	1980

- 172 -

CCCCGGTCA	GAGCCAGCAA GGGGACGGTT CACGGTAAGG TCTGAGGGAG AGAGAGCTCC	2040
TGAGAAACTT	GGGGGGCGCG ACACAGATAG GGTGAAAGCA GAGTGATAGA CCTGGGATGG	2100
TTAGGGGACC	AAGGGAAGAC CAGGCTGGTT GGCATAACACC GGTGAACGGA TGGGAGTCCT	2160
AGGGAAAGAT	GATGCCCTA ACAGTCCTTT CTGTCTCCAC ACCACTCCAG GGGACGATCC	2220
GGAGCTCAAC	TTTCAAAAGC GAGACGCCCC AGCAAGCCTG TTTTGAGAAG TTCTTCAGCG	2280
GCTCTCCTCA	TGGGCCAGAC GGCCCTGGCA AGGGGCAGCA GCAGCACCCC TACCTCGCAG	2340
GCTCTGTACT	CGGACTTCTC TCCTCCGAG GGCTTGGAGG AGCTCCTGTC TGCTCCCCCT	2400
CCTGACCTGG	TTGCCCAACG GCACCAACGGC TGGAACCCCCA AGGATTGCTC CGAGAACATC	2460
GATGTCAAGG	AAGGGGGTCT GTGCTTGAG CGGCGCCCTG TGGCCAGAG CACTGATGGA	2520
GTCCGGGGGA	AACGGGGCTA TTCGAGAGGT CTGCACGCCT GGGAGATCAG CTGGCCCTG	2580
GAGCAAAGGG	GCACACACGC CGTGGTGGC GTGGCCACCG CCCTCGCCCC GCTGCAGGCT	2640
GACCACTATG	CGGCGCTTTT GGGCAGCAAC AGCGAGTCCT GGGGCTGGGA TATGGGCGG	2700
GGAAAATTGT	ATCATCAGAG TAAGGGCCTC GAGGCCCCC AGTATCCAGC TGGACCTCAG	2760
GGTGAGCAGC	TAGTGGTGCC AGAGAGACTG CTGGTGGTTC TGGACATGGA GGAGGGGACT	2820
CTTGGCTACT	CTATTGGGGG CACGTACCTG GGACCAGCCT TCCGTGGACT GAAGGGGAGG	2880
ACCCCTATC	CCTCTGTAAAG TGCTTTGG GGCCAGTGCC AGGTCCGCAT CCGCTACATG	2940
GGCGAAAGAA	GAGGTGAGAT ACGGACTAGG TGTGGGGAGA TCACTACTCT TGGCAATGGT	3000
TTGGGCTGGA	AACTCATGGT TGGAGCACAG GAAGTAGGCT TCTTGTCACT TTGGCCTGTC	3060
ACTTAGATGG	CCTTGGATCT AGCTTCACTC CCAATCCCTA TTGGATGTGA TGCACAAATT	3120
CAGAGCCTTT	GGGTCTCCCT CAGCTGAGGT GGCGGTGGAA ATGGAGGAAG AAGGAAGGGT	3180
GCCTGAGCAG	GATCTCAAGT TCAAGGATGC CTGGAGTTGC TTACTTACCT TGTCTTCCTT	3240
CTCTCTCCGC	AGTGGAGGAA CCACAATCCC TTCTGCACCT GAGCCGCCTG TGTGTGCGCC	3300
ATGCTCTGGG	GGACACCCGG CTGGGTCAAA TATCCACTCT GCCTTTGCCS CCTGCCATGA	3360

- 173 -

AGCGCTATCT GCTCTACAAA TGACCCAGTA GTACACGGTG TGCTGGCACC CTACCGTGGG	3420
GACAGGTGGA GAGGCACCCG CTGGCCTAGA CAACTTTAAA AAGCTGGTGA AGCTGGGGGG	3480
GGGGGGCTGG ACCCCCCCAC CTCCCCCTCT CACAGGAGCA AGACATATAG AAATGATATT	3540
AAACACCATG GCAGCCTGGG ACAAAAGAGGT TTTTGAAGTA AAAAATGAGA TGTATTGTCA	3600
CAACCTGTTT CATTATTGTT TTTTGTGTTG TTTTACACTC CCCCACCCCA GGCTAGAGCC	3660
CCATCACTGT CTTAAGGAAT TATGACAACC CACAAAGCTC AGGCCAGGT GTTATTTC	3720
CTTACATGTA GGATGGTTCA CAAACACAAT ACAGGGGCTT TGGCACCGTG GGGGAGGGGA	3780
CTATCCCAGG CCTCTTAGGG TCTCATGTAT ACCGAATTCA GACCCGAAAG CTCTGAATT	3840
CTGCATCAGA CATCCAGTAG AACTTGGGAG TGAAGCTAGA GCCAAGGCCA TCTAAGTGAC	3900
AGGCCAAAGT GACACGAAGC CCACCTCCTG TGCTCCAACC ATGAGTTCC AGCCCAAACC	3960
AATGGAAGGT GATTCACTT GTCAGGGCCC AAAGGGACAG TCAGTTCTAC TCCCTCCCCT	4020
CACTAGGAGC CACCTTGGTG ACAGTTGATT CTACCCACTG TAAGTGGTAA AGGGATTGGC	4080
CTGGTCCCAA CCATAATAGG CGGGTGGAAA CGGCTCAGGA GGGTACAGCG TGGATTAGGC	4140
CACAAGATGG GGCAGATGAT GTCATCAGAA GCATGTGACC GGTGGGAGCA GTTACTAAC	4200
TTCTGGCAA CCTAGTCCCAT GCTATGCAGG CAGGTAGAGG GATGGGCAGT GCTCATTGTT	4260
TGGCATTGAT GATGTCCACA AATTCAAGGCT TGAGAGATGC GCCACCCACA AGGAAGCCGT	4320
CCACGTCAGG CTGGCTTGCC AGCTCTTGC AGGTTGCTCC AGTCACAGAA CCTGTACCA	4380
GAACAAGAAG ACAGTTGGT CAGGTCTATG ATCAGAACAC TTAAGCCCCA CCTCTCTGTG	4440
CAAGGCAGCC TCAGTCTGTC TTAGCCCATT TCCGTCTTAG CTAGAGCCAA AGCCACTCAC	4500
CTCCATAAAT GATCCGGGTG CTCTGAGCCA CCCCACATCATT GACATTGGAT TTCAGCCATC	4560
CCCGGAGCTT CTCGTGTACT TCCTGTGCCT AGAAGGAGGA GGCAGAGCTA CTAAGTAAGC	4620
TCCCTCCTAT CTATCATTCA AGGAGTAAAA ACCACTGGTT CTCACATAGA GTTGAGTTTC	4680
CAGAAAAGCC CCGGGACCAC AGAGTGGCAA GGCTCCAATC CCACCAAGGCT TGGAATGAAC	4740

08082550 - 4034107

G1

ATTTTGGCA	AAGTCACTCT	CCTTGGTGAG	TTTGGGGGCC	CTCTGTCTCT	AAAGGGGCTT	4800
GGATGGGCTC	CATAGCTGTG	TGAGTCCTGTT	AAAGCCGGAC	AGGCTGAGGA	GCTCTGGGTA	4860
GTTACCTGCT	GAGGGGTTGC	CGTCTTGCCA	GTCCCAATGG	CCCACACAGG	TTCATAGGCC	4920
AGGACCACCT	TGCTCCAGTC	TTTCACATTA	TCTGTGGGGC	AGAGAGGAGA	GTGAGTAGGA	4980
AGGAGCTGAC	CCGCCAAGC					4999

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met Gly Gln Thr Ala Leu Ala Arg Gly Ser Ser Ser Thr Pro Thr Ser
 1 5 10 15

Gln Ala Leu Tyr Ser Asp Phe Ser Pro Pro Glu Gly Leu Glu Glu Leu
— 20 25 30

Leu Ser Ala Pro Pro Pro Asp Leu Val Ala Gln Arg His His Gly Trp
35 40 45

Asn Pro Lys Asp Cys Ser Glu Asn Ile Asp Val Lys Glu Gly Gly Leu
50 55 60

Cys Phe Glu Arg | Arg Pro Val Ala Gln Ser Thr Asp Gly Val Arg Gly
 65 70 75 80

Lys Arg Gly Tyr Ser Arg Gly Leu His Ala Trp Glu Ile Ser Trp Pro
 85 90 95

Leu Glu Gln Arg Gly Thr His Ala Val Val Gly Val Ala Thr Ala Leu
100 105 110

- 175 -

Ala Pro Leu Gln Ala Asp His Tyr Ala Ala Leu Leu Gly Ser Asn Ser			
115	120	125	
Glu Ser Trp Gly Trp Asp Ile Gly Arg Gly Lys Leu Tyr His Gln Ser			
130	135	140	
Lys Gly Leu Glu Ala Pro Gln Tyr Pro Ala Gly Pro Gln Gly Glu Gln			
145	150	155	160
Leu Val Val Pro Glu Arg Leu Leu Val Val Leu Asp Met Glu Glu Gly			
165	170	175	
Thr Leu Gly Tyr Ser Ile Gly Gly Thr Tyr Leu Gly Pro Ala Phe Arg			
180	185	190	
Gly Leu Lys Gly Arg Thr Leu Tyr Pro Ser Val Ser Ala Val Trp Gly			
195	200	205	
Gln Cys Gln Val Arg Ile Arg Tyr Met Gly Glu Arg Arg Val Glu Glu			
210	215	220	
Pro Gln Ser Ieu Leu His Leu Ser Arg Leu Cys Val Arg His Ala Leu			
225	230	235	240
Gly Asp Thr Arg Leu Gly Gln Ile Ser Thr Leu Pro Leu Pro Pro Ala			
245	250	255	
Met Lys Arg Tyr Leu Leu Tyr Lys			
—	260		

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5615 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTACTTTCTT TATATCTCCA TAATTTATT TACTATTACT ACATGATACA TTATTTATA	60
AAAGTCTTTC TAACCTCCTT AAGGATTCAC TGCTTAATCT CCAGTGCTTA GCACAAATCA	120
TTAAATGCGA ACCAGAAACT CTTCCAAATG TGTTACATCT ATAACCTCAT TGGATTCTCA	180
CTACCAACCC CATGCAATAG ATACTAATGT GATCTCTGTC TTACAGAGGA AGAACACAGGC	240
ACAGGGAGGT TCAGTAATTT GCCCAAGGTC ATACACACAC TGGCCTTCAG GTATTCACTGC	300
CCGGGGAGTC TGGTCCCACA GCTGGCATGT TTGCCATTAT ATTATATTGC CTCCTTATAG	360
TGTCGGCACT CATTAAAGCAC ATTGACAGCT ATGCTTGGTG AGTGAECTACT ATGTACCCAG	420
CTCTGTGCTA CATGCTTAC CTGGATTATT TCAACTGCAC AACAAACCTG TGAGGTAAC	480
ACCATCATTG CTCCTATTT ACATAACAGA AAACTACAGA AATCTGGGC TGGGGTAGT	540
GGCTCATGCC TGAAATCCA GCACCTTGGG AGACCCGTGC TCTAAAAAAA ATTTTTTTT	600
GGCCGGACGT GGTGGCTCAC ACCTGTAATC TCAGCACTTT GGGAGGCTAA GGCAGGCAGA	660
TCACAAGGTC AGGAGTTCTA GACCAGCCTG GCCAACATGG CAAAACCTG TGTCTACTAA	720
AAATACAAAA AATAGCTAGG CGTGGTGGCA GGTGCCTGTA ATCCCAGCTA CTCAGGAGGC	780
TGAGGCAGGA GAATCCCCCTG AACCTGGGAG ATGGAGGTTA CAGAGAGCCG AGATCGTGCC	840
GCTGCACCTCC AGCCTGGCCA ACAAGAGCAA GACTCTGTCT CGAAAAAAAT AAAAATAAAA	900
ATAAAAATAT TTTTTAAAAA ATTAGCTGGG TGTGGTAGCA CATGCCTGTA GTCCCAGCTA	960
CTTGGGAGGC TGAGGTAGGA CGATCACTTG AGCCCAGGAG GTCAAGGCTG CAGTGGCTG	1020
TGATGGGCC ACTGCACCTCT AGCCTTGGTG ACAGCAAGAC CCTGTCTCAA AAAAAAAA	1080
AAGAGAAATC GGGCAACTTC CCCAAGATCG CGCAGTTAAC TAGTGGCATA GCTTCACTCA	1140
AACTCGAAGT CTTAACAGG AACTCTACC AAATGAGATC AACGGCTCAG TAATGGATTG	1200
GCATCCAGTA TGAAGACTGG ACCAGCAGGG AGAACTATGA TGCGTACAGC CTAGAGCCTG	1260
AAGCAGATTG CACAGCCTCA GAGGTGGCAC AGGCTGACTC ACAACCCGGG GCAGAAAGGG	1320
ACCAGCCAG AAACAGTGAC CCAGAACATCAC AGGGAAGTAG AAATGGGATT CGGCACAAATG	1380

- 177 -

AAGCCCCCTCC	TTGACCCCAT	GTCCTTACC	CTCAGGGGGC	CAGGAGTTAG	TCGCTCAGGC	1440
GGCTCAAAGG	TCTTGACGGT	GGAGAACACC	ATCCCCAGGG	ATTCCCGACG	CGGTGATGCC	1500
ATCAAAGCGT	TAATTCTGAG	ATGGGCCTGC	CCGGGTGCGG	ACTCTGCCGC	AGCAAGAGAA	1560
GGGTTAACTG	CCCCGGGCCT	TCGCCGTGGG	GGCGGGGCCT	CGGGGAGGGT	CACAGCCC GG	1620
GAATGAGACC	CGAGGTTAAC	CGCCCGGGGT	GGGCTCCACG	GGGGCGGGGC	ATGCTCTCCG	1680
CGGCTGCTGC	CGGTATAAGAG	CGGTAAC TGC	CCAGGAGGGG	GGGGGGCCCC	ACAGGGGGCT	1740
GGCCTCGGAG	CTGCACGGCC	GTGGGC GGCG	ATGAGAGGGT	TAAGCCCCAG	AGGGCCCTGG	1800
AGGGGCGGGG	CCGC GGGGACG	GGCTCGGCC	AAGGGAGGAG	CTGGGGCGG	AA CGGGCCGG	1860
CGGTCTGC	CCTGC GGC	TCGGCTTCTT	TCCGCCGGC	TCCTTCAGAG	GCCCGGGCAC	1920
CTCCAGGGCT	GGGAAGTCAA	CCGAGGTTCG	GGGCAGCGG	CGAGGGCTCC	GGGCAGTAA	1980
GGGGGATGGT	CCATGCTGAG	GCCCCAATGG	GGCGAACTCG	CGAGAGTCTC	TGGCGACCTG	2040
GATCAGATGG	GGCGAGGGCA	GATGAAGGGC	CCAGGAGCTT	TGGGGCAGCG	AGGAGGGAGG	2100
AGCGGGCCCG	TTGGCAA ACT	TGGGTGAAAG	CATGGGGTAC	CTGGGTGACG	AGCCCCGCC	2160
AGGATTCTGC	TCTTCACGCC	CCTTTCTCC	CAGCTCCCTT	CCAGGTCAAT	CCAAACTGG	2220
GCTCAAGTTT	CAGAAGAGAA	AGACGCCCA	GCAAGCCTCT	TTCGGGGAGT	CCTCTAGCTC	2280
CTCACCTCCA	TGGGCCAGAC	AGCTCTGGCA	GGGGCAGCA	GCAGCACCCC	CACGCCACAG	2340
GCCCTGTACC	CTGACCTCTC	CTGTCCCAG	GGCTTGGAAAG	AGCTGCTGTC	TGCACCCCT	2400
CCTGACCTGG	GGGCC CAGCG	GCGCCACGGT	TGGAACCCCA	AAGACTGTTC	AGAGAACATC	2460
GAGGTCAAGG	AAGGAGGGTT	GTACTTTGAG	CGCGGGCCCG	TGGCCAGAG	CACTGATGGG	2520
GCCCCGGGTA	AGAGGGGCTA	TTCAAGGGC	CTGCACGCC	GGGAGATCAG	CTGGCCCCTA	2580
GAGCAGAGGG	GCACGCATGC	CGTGGTGGG	GTGGCCACGG	CCCTCGCCCC	GCTGCAGACT	2640
GACCACTACG	CGGC GCTGCT	GGGCAGCAAC	AGCGAGTCGT	GGGGCTGGGA	CATCGGGCGG	2700
GGGAAGCTGT	ACCATCAGAG	CAAGGGCCCC	GGAGCCCCCC	AGTATCCAGC	GGGAAC TCA	2760

- 178 -

GGTGAGCAGC TCGAGGTGCC AGAGAGACTG CTGGTGGTTC TGGACATGGA GGAGGGAACT	2820
CTGGGCTACG CTATTGGGG CACCTACCTG GGGCCAGCAT TCCGCGGACT GAAGGGCAGG	2880
ACCCCTCTATC CGGCAGTAAG CGCTGTCTGG GGCCAGTGCC AGGTCCGCAT CCGCTACCTG	2940
GGCGAAAGGA GAGGTGAGGC CTGGGGCAGA CGTGGGGAGA ACTTTCTGTC CCTGGTGGCA	3000
GTGGTTTGGG ATGGAAACTC TTCTGACAAG AGCAGAGGGG ATGGACCTTC ATCCAGCCTG	3060
CCTCAACCTC TGTTCAAGTGC TGGGAAAGGC TAGGGGTCTT CACAGCTGTT ATTTAATTAA	3120
ACCCAACAGC AATAGAGGTG AAACAGGCTT GAGAAAGCAA CTTTCTCAAG TTCTCTTGGC	3180
CACTAAATGG TGAACCTTCA GAATGGAGGG AGGAACGTGCA GGGATGAGAG AATTCAAGGAG	3240
ATATCAACCC CTGAGCAAGA GGTGCAAAGC GTTAGGTACT GGGTTTGATG TACAGGTCCA	3300
AAAGAAGGAT GGGCAGAGGC AGGTACCCAG GCTGTATAACC GGATTCCCTG GGCTCTAACCC	3360
TGTCTCTGTG CCACATACCT ACTTCCTTCC TCAGCCACAC CTCTGGATGG AGACACTGGG	3420
GCCCTGGCA CCAGGGAGGA GAGCACTGGA GGAGGCAGGG CCTTAGGGTG GGGCAGCAGG	3480
GGAGGAGCCT CCCCAGGAAC TGACTGGTC CAGGGCTTGG AGCTGCTCTC TGCAGTTGTG	3540
TGGGCTGTAG AGTGGAGGGC CATCCCTCCT CACCTCAGCC CCAGCTCCCA AGCCTCTGGA	3600
GTCAAAAGCCT GGGCCAGCTC CACCACTGTC AGAGCCACCT TGGCCTGTTG TTTAGAGGGC	3660
CTTAGCCAGC TCTTCACCCC CAGCTCTGAC TAGGGATGTG TGAAATCTTA TCTGGGAGGC	3720
AGAACTTCCG GGTATCTCAA ATTCCCCTTT CAGCCAGGTG GGCACACTCG AAGCAGGAAA	3780
GCAGAAAGGC ATCTGAGTAG GACCCCGTAG TTTGAGGACA TCTGGCTGGT GGCTGCACCC	3840
ATACTTACAT TCCCCTCCCTT CTCTCTCCCA GCGGAGGCCAC ACTCCCTTCT GCACCTGAGC	3900
CGCCTGTGTG TGCGCCACAA CCTGGGGGAT ACCCGGCTCG GCCAGGTGTC TGCCCTGCC	3960
TTGCCCCCTG CCATGAAGCG CTACCTGCTC TACCAAGTGAG CCCTGTGATA CCACAGACTG	4020
TGCTGAGGTC TTGCCACAC CCCTCCCCTT GGGGAGGTGG GGAGGCAGTG CTGGCCTAGA	4080
CCAGCTGCTG AAAGCTGGTG AGGCTGAGCC CCTACCCCAA CCCAAGCTCT GCGGAAATCA	4140

ACAGCCCCAG ACCCACTTGG AGGGAGGAAG AAAGGGAGCC GGCCTCAAG GCTATGACAG	4200
TCTGCTACGC AAAACATTTT TTCAAGTAAA AATAGTAAGA GATGTTGTTA TAGAAACCTG	4260
TTCTTGTTTT TTTTTTTTC TTGCACAAAT GATCATTTAT ATAGCTGCCT CAAAAAGGAA	4320
GATTATCTGG GCAAGTCCAG TGAAGGCAGA CAAACCACAA GACCTAGTGC CAGGTTTATT	4380
CCCTCACATG GGTGGTTCAC ATACACAGCA CAGAGGCACG GGCACCATGG GAGAGGGCAG	4440
CACTCCTGCC TTCTGAGGGG ATCTTGGCCT CACGGTGTAA GAAGGGAGAG GATGGTTCT	4500
CTTCTGCCCT CACTAGGGCC TAGGGAACCC AGGAGCAAAT CCCACCAACGC CTTCCATCTC	4560
TCAGCCAAGG AGAACCCACC TTGGTGACGT TTAGTTCCAA CCATTATACT AAGTGGAGAA	4620
GGGATTGGCC TGGTCCCAAC CATTACAGGG TGAAGATATA AACAGTAAAG GAAGATAACAG	4680
TTTGGATGAG GCCACAGGAA GGACCGAGATG ACACCATCAG AAGCATATGC AGGGAAAGGG	4740
CAGTTACTGG GCTTCCTGGC TGCTTAGTCC CTGGCTTGGC AGGAAGGGTA GGGAAAGATGG	4800
ATGGGGCTCA TTGTTTGCA TTGATGATGT CCACGAATTG GGGCTTGAGG GAAGCACCAC	4860
CCACAAGGAA GCCATCCACA TCAGGCTGGC TGGCCAGCTC CTTGCAGGTT GCCCCAGTCA	4920
CAGAGCCTGG GAAGGGAGCA GAACAAGGGC TTGGTCAAGA ATGGGATGAG TCTGCCCAT	4980
CCCCAGGTCC ATGTCCGAGG GCTCAGTCTA GTCTCAGCC CACTCCACCT CAGCCGGAA	5040
CCAAAGCCAC TCACCTCCAT AAATGATACG GGTGCTCTGA GCCACCGCAT CAGAGACGTT	5100
GGACTTCAGC CATCCTCGGA GCTTCTCGTG TACTTCCTGG GCCTAGAACCA AGAAGCTGGC	5160
CTAAGTAAGA CCTTTCTGCA CTCTCTAAGA GGAAAAATCA CTGGCACCAAG TGGACACTTA	5220
GTGTGGTTTC TGACTGAGTC AGAGTACCAAG GGCTCTGATC CAAGCCAGGC CCTGGACTGG	5280
ATGCCCTTGG ACAAGTCAGT GTCTCTGGT TCAAGGTCTC TGTGTCTTG AAATAAGGGG	5340
TGCCCCATG TGGGCTGTGT CTGTCCAAAC CTATTGAGGC AGGCTGGGAT CACGGCAGGG	5400
CTCCTGGGCC CGGTTACCTG TTGGGGTGTG GCAGTCTTGC CAGTACCAAT GGCCCACACA	5460
GGCTCATAGG CCAGGACGAC CTTGCTCCAG TCCTTCACGT TATCTGCAGG GCAGAGATAC	5520

G1

AGATGGAGGG AAGGGTGAAC AAGAAAGAGC TCTCCAGCCA GGTTCTCCGG AGTACGAAGA 5580
ACGGTGGCCT ACTGCCCCCT AGTGGACATT GGGGG 5615

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Gln Ala Leu Tyr Pro Asp Leu Ser Cys Pro Glu Gly Leu Glu Glu Leu
 20 25 30

Leu Ser Ala Pro Pro Pro Asp Leu Gly Ala Gln Arg Arg His Gly Trp
35 40 45

Asn	Pro	Lys	Asp	Cys	Ser	Glu	Asn	Ile	Glu	Val	Lys	Glu	Gly	Gly	Leu
50							55					60			

Tyr Phe Glu Arg Arg Pro Val Ala Gln Ser Thr Asp Gly Ala Arg Gly
 65 70 75 80

Lys Arg Gly Tyr Ser Arg Gly Leu His Ala Trp Glu Ile Ser Trp Pro
85 90 95

Leu Glu Gln Arg Gly Thr His Ala Val Val Gly Val Ala Thr Ala Leu
100 105 110

Ala Pro Leu Gln Thr Asp His Tyr Ala Ala Leu Gly Ser Asn Ser
115 120 125

Glu Ser Trp Gly Trp Asp Ile Gly Arg Gly Lys Leu Tyr His Gln Ser

- 181 -

130	135	140	
Lys	Gly Pro Gly Ala Pro Gln Tyr Pro Ala Gly Thr Gln Gly Glu Gln		
145	150	155	160
Leu	Glu Val Pro Glu Arg Leu Leu Val Val Leu Asp Met Glu Glu Gly		
	165	170	175
Thr	Leu Gly Tyr Ala Ile Gly Gly Thr Tyr Leu Gly Pro Ala Phe Arg		
	180	185	190
Gly	Leu Lys Gly Arg Thr Leu Tyr Pro Ala Val Ser Ala Val Trp Gly		
	195	200	205
Gln	Cys Gln Val Arg Ile Arg Tyr Leu Gly Glu Arg Arg Ala Glu Pro		
	210	215	220
His	Ser Leu Leu His Leu Ser Arg Leu Cys Val Arg His Asn Leu Gly		
225	230	235	240
Asp	Thr Arg Leu Gly Gln Val Ser Ala Leu Pro Leu Pro Pro Ala Met		
	245	250	255
Lys	Arg Tyr Leu Leu Tyr Gln		
	260		

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGCTAGATCTGGACCCCTACAATGGCAGC

28

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: base pairs

- 182 -

- (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

G1
(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGCTAGATCT GCCATCCTAC TCGAGGGGCC AGCTGG

36